PERIODICAL ROOM GENERAL LIBRARY UNIV. OF MICH.
MAR 17 1926
cop 2

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS

HAROLD RUGG, *Chairman.*

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN
FRANK N. FREEMAN — V. A. C. HENMON — BEARDSLEY RUML — E. L. THORNDIKE

VOL. XVII. No. 3. MARCH, 1926 $4.00 a year. 60c per copy.

CONTENTS

FIRST STEPS TOWARD A SCALE FOR MEASURING ATTITUDES. *Mark A. May and Hugh Hartshorne* 145

A PERSONAL CONSTANT. *H. Heinis* 163

A TEST FOR ART APPRECIATION. *Theodore F. Karwoski and Erwin O. Christensen* 187

AGE, SEX AND THE INTERRELATIONS OF MENTAL ATTAINMENTS OF COLLEGE STUDENTS. *John W. Gowen and Marjorie Gooch* 195

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 208

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 211

Published Monthly Except June to August by

Warwick & York, Inc.,

York, Pa. Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. Title page and index are bound in December number.

Manuscripts for publication, books, or other materials for review, and news items should be addressed to Harold Rugg, Chairman of the Editorial Board, 400 W. 118th Street, New York City.

The price of the Journal is $4.00 a year, payable in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for non-receipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,
Baltimore, Md.

University of Virginia

Summer Quarter

Edwin Anderson Alderman, Ph.B., D.C.L., LL.D., President
Charles Gilmore Maphis, Ped.D., LL.D., Dean

First Term, June 21st-July 31st
Second Term, August 2nd-September 4th

The Summer Quarter is an integral part of the University year, the courses being the same in character and credit value as in the other quarters of the year.

Degrees are conferred upon men and women for summer work.

The Master's Degree may be obtained by properly qualified students in three summer quarters.

It offers opportunities unexcelled in the South and makes a strong appeal to teachers seeking broader scholarship and training and wider social contacts, and to college students desiring to complete degree requirements.

Attendance last quarter, 2212 from thirty-nine states and foreign countries.

The most beautiful and unique campus in America.

Accommodations at reasonable rates. Tuition for non-Virginia students, $20 per term.

Entertainments, music festival, excursions.

Reduced railroad rates.

For illustrated folder and full announcement write to

Secretary, Summer Quarter, University, Virginia

The "Supreme Authority"

WEBSTER'S NEW INTERNATIONAL DICTIONARY

—THE MERRIAM-WEBSTER

Because

Hundreds of Supreme Court Judges concur in highest praise of the work as their *Authority.*

The Presidents of all leading Universities Colleges, and Normal Schools give their hearty *indorsement.*

All States that have adopted a large dictionary as *standard* have selected Webster's New International.

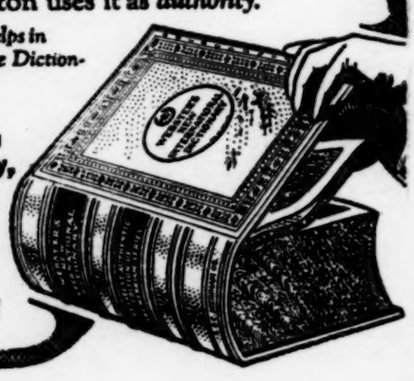
The Schoolbooks of the Country *adhere* to the Merriam-Webster system of diacritical marks.

The Government Printing Office at Washington uses it as *authority.*

Write for Helps in Teaching the Dictionary, FREE.

G. & C. Merriam Company,
Springfield, Mass.

Get The Best!



TEACHERS' HEADQUARTERS

European Plan

Entirely Fire-Proof

Centrally Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XVII March, 1926 Number 3

FIRST STEPS TOWARD A SCALE FOR MEASURING ATTITUDES

MARK A. MAY AND HUGH HARTSHORNE[1]

Teachers College, Columbia University

We are accustomed to measuring physical energy by the amount of work done against standardized resistances, or standardized ways of measuring resistance. If we are measuring power, we may take gravity as the fundamental constant. Having a unit of weight called a pound, a unit of distance called a foot and a unit of time called a minute, we can then state a unit of energy in terms of the amount needed to raise a pound a foot in a minute.

If we are measuring heat, we may take the caloric inertia of a common substance, like water, as the fundamental constant. Having already a scale for measuring the intensity of heat, and units of measure for determining amounts of water, we can state the amount of heat in terms of what is needed to raise a unit of water a degree, or unit of heat intensity.

In these and similar cases measurement of amount of energy has depended on previously determined space, weight, and time units, in terms of which the work done is measured. When we set out to measure mental energy, however, we find no such previously determined units of resistance in terms of which the work done may be measured. Since psychology has no such units it must resort to other devices that will take their place for practical work.

[1] Dr. May and Dr. Hartshorne are the investigators for the Character Education Inquiry, which is being conducted by the Institute of Educational Research, Teachers College, in cooperation with the Institute of Social and Religious Research, New York.

The concepts of both "resistance" and a "unit," as used commonly in the making of educational and psychological tests, have been developed out of the characteristics of the probability curve. For "resistance" is substituted "difficulty." To spell a hard word is to overcome a greater "resistance" so to speak, than to spell an easy one. The order of difficulty is usually established by the varying percentages of group performance or by order of merit ratings of competent judges. A task that can be done by only a few is hard. A task that only a few cannot do is easy. A scale is thus established by the administration of the tasks to an unselected population in which the trait measured is presumably distributed on a normal frequency surface. Position on the scale thus measures the amount of difficulty overcome, and therefore, the "energy" or ability used to overcome it.

The *unit* of measurement, where units are used, is the standard deviation, or some other measure of variability of a normal curve, by which a section of the frequency surface can be translated into distances along the base line. The zero point is either an arbitrary point of reference, such as the mean of the group, or plus or minus 3SD, or an absolute zero established experimentally. Experimentation has revealed the fact that many abilities measured in this way increase with age. This has suggested another way of defining difficulty, *viz.*, in terms of what a certain proportion of persons of a certain age can do. Thus has arisen the concept of mental age and the mental age scale.

Since 1908, when Binet published the mental age scale, a vast amount of work has been done in applying these fundamental concepts to the measurement of mental abilities of various sorts. The work has centered around general intelligence on the one hand, and school room activities on the other, with more recent excursions into the field of vocational activities. Arithmetic, spelling, vocabulary, language, silent reading, etc., and, with different scale ideas, hand-writing and composition, represent the sort of abilities that have been tested and scaled.

The scales that have been developed have this added characteristic that, beginning at the bottom, any point on the scale includes all points beneath it. The individual tested may strike in at any point on the scale with some assurance that if he can perform the task assigned for that point he can perform all the tasks below it. In actual practice, of course, there is considerable scattering, but it is remarkable to what an extent the rule holds true.

Does the same fundamental notion of a scale, no matter how determined, apply to other sorts of behavior tendencies, such as attitudes? There have been sporadic efforts to measure these other tendencies, and now and then an attempt to apply the scale idea. The Character Education Inquiry has been fortunate in its control of facilities which have enabled it to experiment somewhat extensively in this field. A portion of one such study, looking toward the building of a scale for the measurement of an attitude, will be discussed in this article.

The particular attitude or behavior tendency selected is dishonesty. For the purposes of measurement, we may assume that dishonesty is analogous to a particular form of energy, such as heat, and that it is to be measured by the amount of work done, that is, by the dishonest acts performed. Reverting again to the analogy of physical energy, the work done is to be measured by the resistance overcome in doing it. As in the case of general intelligence, we do not have ready to hand any previously standardized units of measurement by which we can define our resistances. We shall therefore have to devise them. And from past experience let us take the cue of measuring difficulty, or resistance, by reference to the distribution of the tendency to be measured. This will require that we select resistances to dishonsty that are common to the entire population tested, and that we standardize these resistances, always using the same ones.

It may be questioned at this point whether we are justified in assuming the normal distribution of such a tendency as dishonesty and basing a scale on this assumption. For purposes of experiment we believe we are and that only by experiment can the need of modifications in the assumption be determined.

Building a Scale for Measuring Dishonesty

By virtue of past experiences, or heredity, or both, the individual may be more or less permanently predisposed to make an honest or dishonest response in situations involving this type of conduct. Here is a man faced with a mass of figures representing his income for the year of 1925. If his neural mechanism is set or predisposed to facing the facts regardless of cost or consequence, he will be more likely to make an honest return than he would be if his neurones were set in the direction of shading the truth or "doctoring" figures so that the result will be favorable to himself. His behavior, however, is not totally determined by his attitude on the question of honesty. Other

powerful but more temporary sets may operate. For the purpose of further discussion let us call the permanent sets "attitudes" and the temporary sets "desires." In order to test the attitudes, the desires must be kept constant. Accordingly the test scale must consist of a series of situations in which behavior is demotivated as far as possible.

With these conditions in mind, let us endeavor to construct a scale for measuring the attitudes of school children toward cheating on examinations. Assume that by the time a child reaches the fifth grade he has more or less permanent sets or attitudes toward cheating in school work. He may be set against it, or be in favor of it, or be neutral toward it.[1]

We are assuming that these sets are variables and that they run from neutrality to all degrees of *pro-ness* in one direction and in the opposite direction to all degrees of *anti-ness*. At one end are those who cannot help cheating. At the other end are those who are "dead set" against it. We are proposing, then, to test the varying degrees of *pro* attitude by the resistance overcome in order to cheat and the degrees of *anti* attitude by the strength of the motive it takes to induce cheating. The first step is to set up a kind of middle situation which will separate the *pros* from the *antis*. This situation should be so arranged that it will be clear to the child that nothing will be gained or lost by cheating, and that if he cheats no one will know it and if he is honest no one will know it. Cheating here should be relatively easy. The situation should then contain these elements.

(*a*) Cheating is very easy, in fact so easy that it requires less mental effort to cheat than to be honest.
(*b*) There is nothing to be gained from cheating.
(*c*) There is nothing to be lost by it.
(*d*) There is no apparent chance of any one's knowing whether cheating has taken place or not.

It is assumed that those who cheat under these conditions have a tendency toward it, and that those who do not cheat have a tendency against it or else are neutral. From this middle point, therefore, the scale is constructed. By keeping all aspects of this general situation constant except the first element, that of difficulty of cheating, the

[1] It may be true that no child is set directly in favor of cheating as an end in itself. The set or predisposition may always include a drive toward the thing that cheating will secure. In any case his attitudes are such as to favor a cheating response.

scale roughly measures pro-cheating attitude. By keeping the entire situation constant and by adding on a series of constant motives, the scale measures the anti-cheating attitude.

One of the scales with which we have been working consists of seven steps, five *pro* and two *anti*. Beginning at the extreme of the *pro* attitude the steps are defined by the behavior of the children on the following tests.

1. An Information test[1] consisting of 28 items steeply graded in difficulty. Instead of underlining the correct answer, the pupil is required to encircle it in *ink*. He is told to guess if he does not know. In fact, he is not allowed to hand in his paper until he has guessed at every answer. These papers are taken to the office where a duplicate of each is made. A day or so later the original papers are returned to the children with answer sheets and they are instructed to grade their own papers. Each child has previously been supplied with an ink eraser in connection with his school work.

In order to cheat on this test it is necessary for the child to erase a circle drawn in ink and make another. This is rather difficult. It is not easy to make a clean job of it. By comparing the corrected paper with its duplicate in the office, it is possible to see how much cheating has gone on.

2. An original Disarranged Sentence test requiring the use of pencil. The task is to straighten out mixed up words so they will make a sentence. In case the child finds one too difficult, and they are sharply scaled in difficulty, he is required to draw a line through the blank space to indicate that he has definitely given it up and omitted it. When the papers are redistributed with answer sheets, cheating is accomplished by erasing a whole sentence written in pencil. This is easier to do than to erase a circle made in ink. Consequently a greater percentage cheat on Test 2 than on Test 1.

3. The Thorndike-McCall Reading Scale. The same procedure is used here as for Test 2. Cheating is done by changing one or two words—sometimes a whole phrase—and by adding answers. Pencils are used.

4. A Sentence Completion test. Cheating here consists in either adding on more words, that is, doing more items, or in changing the words previously written, by erasing and rewriting.

5. An Original Spelling test. It is made up of 90 words some of which are misspelled. The task is to check the misspelled words. This test represents the *very easy* level of cheating. All the child has to do is to add more check marks or erase those previously made. It is very easy to erase a check mark. This test is intended to strike in at about the neutral level outlined above. Cheating on it is easy and as far as we know no child has suspected that we keep check on the amount of cheating. It is not directly motivated although there is, of course, such habitual motivation as is involved in routine school work.

6. A Word-knowledge test, given with slight motivation, and therefore on the *anti* side of the neutral point. This is arranged as a multiple choice test.

[1] The material for Tests 1, 4, 6 and 7 was furnished by the Institute of Educational Research, Teachers College.

The response words to each stimulus word are numbered one to five. The task is to find the correct response word and write its number on the dotted line in the margin. This makes cheating easier than it would be if the response word were underlined. Motivation is provided by writing the norms for the grade concerned on the blackboard. This sets a goal for each child to work toward.

7. An Arithmetic test, the answers being written at the right margin. The only essential difference between this and Test 6 is that the children are told that this will count in their monthly standing.[1]

To summarize, we have the following series of difficulties on the *pro* end of the scale:

1. To erase a circle in ink and add another.
2. To erase a sentence in pencil and add another.
3. To erase a phrase of two or more words, or add more words.
4. To erase a single word or add a single word.
5. To erase a check mark or add a check mark.

On the *anti* end of the scale:

6. To erase a digit or add a digit. *Motive:* The norms visible on the black board.
7. To erase an arithmetic answer or part of it, or to add an answer. *Motive:* Norms visible on blackboard and an oral statement that this test will count.

Criteria for Evaluating the Scale

The statistical requirements of the proposed scale may be summarized as follows:

1. It should present a series of situations calling for behavior in which cheating is possible and not likely to be discovered.
2. The motives for cheating which are introduced at the *anti* end of the scale, should be connected with tests all of which are neutral in difficulty.
3. The steps on the scale should be approximately equal and in multiples of some unit of measurement.
4. The scale should have as a zero point either an absolute zero or an arbitrary point of reference.
5. The limits of the scale should be, on the *pro* end, a point of difficulty where just no one will cheat or where, say, 1 in a 100,000 will cheat; and at the *anti* end, a situation in which 100 per cent or say, 99,999 in 100,000 will cheat.[2]

---

[1] Some may object that it is improper to motivate cheating or provide an opportunity for it. We have taken extreme precaution to use no motives that are not actually employed in the school concerned and to employ no opportunity that is not similar to opportunities inherent in the school situation. The moral strain under which a child is put on one of these tests is by no means as great as that under which he is put when he is doing work required for promotion, or taking a State examination.

[2] These are, of course, theoretical limits. No one would care to motivate cheating to a point where 100 per cent would cheat even if this should be possible.

TABLE I

	1	2	3	4	5	6	7	8	9		10
	In	DS	R	LC	WK	SP AR	Per cent not skipping	Per cent skipping	Number of skips	Per cent skipping	Per cent of space filled
	6	6	6	6	6	6	100				100
		22	22	22	22	22	55		0	55	82
		5		5	5	5		10	1	27	
		5	5		5	5		10	2	10	
		4	4	4		4		8	3	8	
		4				4		8			
		3			3	3		6			
		1		1		1		2			
		1	1			1		2			
		45									
			58	58	58	58	68.5		0	68.5	90
			12		12	12		13.5	1	27	
			13	13		13		13.5	2	3	
			2			2		2.5	3	1.5	
			2					1.5			
			1	1				1.0			
			88								
				25	25	25	77		0	77	87
				11		11		19	1	20	
				3				3	2	3	
				1	1			1			
				40							
					35	35	98		0	98	98
					3			2	1	2	
					38						
						30					
	6	51	126	150	175	237				Total	89
Per cent cheating	2	17	43	51	59	80					
Per cent scaled cheaters	100 6's	100 5's	86 4's	73 3's	81 2's	96 1's					

6. The situations should be so arranged that all who cheat at one level will also cheat at all lower or easier levels.

7. As a corrolary of six, all those who cheat at least once should cheat on the lowest level, all those who cheat at least twice on the two lowest, and so on.

8. The obstacles should be of such a nature that they can be overcome merely by the application of time and effort, without regard to intelligence, and without the addition of other attitudes.

Experimenting with the Scale

We first tried the scale out on a group of 32 public school pupils. This was a 5*A* group of bright children. Later we tested a larger group of 263 children in an institution for dependents. These were distributed through Grades IV to VIII. The average IQ of this group was 97. Chronological ages ranged from 8 to 16. Eighty-four per cent made at least one change in their papers.

Table I shows the results of the preliminary use of the proposed scale in respect to these 237 children who exhibited the trait in question.

Application of the Criteria

The first criterion has been dealt with in describing the tests. The second criterion is not completely met. The difficulty of the Arithmetic and Word Knowledge tests is not neutral. The results show that about the same percentage cheat on the Spelling test, which was not motivated, as on the Arithmetic test which was. For this reason we combined Arithmetic and Spelling into one step on the scale and placed them at the same level. Thus Table I shows six steps instead of seven. Some idea concerning the relative difficulty which each test situation presents may be had by reference to the mean number of changes made on the test when they were returned to the pupils for scoring. These facts are shown in Table II.

Table II

	Mean	Median
Information	1	
Disarranged Sentences	1.86	.68
Reading	3.40	2.68
Language Completions	7.30	5.24
Word Knowledge	9.87	6.26
Spelling	11.82	9.20
Arithmetic	6.80	5.04

The fact that the Disarranged Sentence test is harder to cheat on than the Reading test is shown not only by the fact that more pupils cheated on the Reading than on the Disarranged Sentences, but that those who cheated made, on an average, 1.86 changes as against an average of 3.4 changes on the other. In spite of the fact that Arithmetic was motivated most strongly of any, the mean number of changes was 6.80 against 11.82 for Spelling, which was not motivated at all.

Common sense would say that it is easier to change a digit in the Word Knowledge test than to change an answer in the Arithmetic test, which usually involves more than one digit, with other symbols. Now if we placed Arithmetic between Completion and Word Knowledge and not ahead of Completion, as the percentage would indicate, and assume that the same number would cheat on it as cheated on the Word Knowledge, then its difficulty becomes 8.4,[1] which places it at this point.

This now gives us a fairly good notion of the order of difficulty of these test situations in regard to cheating and indicates that we have now the beginning of a scale.

The third criterion demands that the steps on the scale be equal and preferably in multiples of some unit of measurement. Theoretically we are attempting to measure attitudes toward cheating in terms of the amount of resistance overcome, but we have no units of resistance. All we can do is to fall back on the normal curve and use the SD as our unit. We do this with our eyes wide open to the dangers and pitfalls involved in assuming a normal distribution of the function we are attempting to test, and a random sampling of cases. Perhaps neither of these assumptions is correct, but as was noted above, needed changes in them can best be determined by experiment.

The usual procedure is, first, to find the per cent who cheated at each level. We get this from the figures at the bottom of Table I. Referring these percentages to the table of probability intervals for the normal curve we assign each an SD value. Thus 2 per cent cheat on the Information test. There are 2 per cent of cases between 2.06 SD and the end of the normal curve. Thus we say that those cheating on the Information Test went beyond the limit of 2.06 SD. Call this plus 2.06 SD, to show the direction from the mean. The SD values for all the steps are given in Table III.

---

[1] Number of changes made on Arithmetic (1313) divided by number who cheated on Word Knowledge (156).

TABLE III

	Per cent cheating	SD	Scale value
Information	2	+2.06	51
Disarranged Sentences	17	+ .96	40
Reading	43	+ .08	31
Language Completion	51	− .28	27
Word Knowledge	59	− .64	24
Arithmetic-spelling	80	−1.00	20

To get rid of the sign we assign arbitrary position values to the sigma steps of the normal curve. Calling plus 3.0 SD zero and proceeding by giving a value of one for each .1 SD we have:

−3.0 SD	−2.5	−2.0	−1.5	−1.0	−.5	0	+.5	+1.0	+1.5	+2.0	+2.5	+3.0
0	5	10	15	20	25	30	35	40	45	50	55	60

This gives to the levels shown in Table III, column two, the values shown in column three. On the basis of the assumptions made above, Table III shows the amount of inequality of the steps.

The fourth criterion requires a zero point either arbitrary or absolute. Assuming that these six steps represent degrees of the *pro* attitude toward cheating then there ought to be a point where the *pro* attitude becomes neutral or swings over to the *anti*. This would be the theoretical absolute zero. If this point could be satisfactorily located and objectively defined, we could mark off the distance between it and that degree of difficulty where just no one will cheat and divide this distance into equal arbitrary units. This would give a scale analogous to the thermometer. In the same way the *anti* attitude would be measured in the other direction. There are two difficulties here. One has to do with defining the neutral attitude,[1] the other is the annoyance of dealing with minus signs—there is some advantage in having zero at the one end of the scale. Our Spelling test, however, is probably near the zero point as defined above, so that if we wished to use the absolute zero or neutral portion as the point of reference we would

[1] For example, actual behavior doubtless involves, in cases near the neutral point, an admixture of both *pro* and *anti* attitudes, the *anti* attitude beginning to operate before the *pro* attitude has reached zero, thus tending to scale off the overt behavior more rapidly than the descending *pro*-attitude alone would account for.

not go far astray if we placed it near the line between the spelling test and those who did not cheat at all; that is, near −1.00 SD or what, on the present scale, is called 20.

Table III shows how well we have met the fifth criterion. Two per cent of our population will cheat on the Information test. Perhaps if we should use a test requiring the erasure of a whole sentence in ink we would approach the limit demanded by the criterion. As to the other extreme, we have already stated that we do not wish to investigate further in this direction.

The sixth criterion for the scale is that when a subject cheats on one level of difficulty he should theoretically cheat on all easier levels. Table I shows the facts in regard to this in great detail. The six who cheated on the Information test cheated also on everything else. But of the 51 who cheated on the Disarranged Sentence test, only 28 cheated on all easier levels, 5 skipped Reading, 5 skipped Completions, 4 skipped Word Knowledge, etc. These facts are summarized in the seventh and eighth columns. In respect to this criterion the steps stand in percentage of efficiency:

Information	100
Disarranged Sentences	55
Reading	68
Completions	77
Word Knowledge	98

It should be noted that apart from the Information test these percentages grow larger as cheating gets easier. One reason for this is that there are four chances to skip below the level of Disarranged Sentences, which is at a high point on the scale. If a pupil cheats on this test he may skip any of the four steps below it. There is more opportunity, hence more skipping. On the other hand those who cheat on the Word Knowledge test have left but one chance to skip. In order to correct for this fact we have computed for each step on the scale the number of opportunities for cheating below it, and used this as a common denominator. For example, on the Disarranged Sentence test there are 51 cheaters, each of whom had 4 opportunities to skip, making a total of 204 opportunities. Now there are actually 36 gaps, which is 18 per cent of 204. Hence of all the possibile gaps, 82 per cent are filled. Our modified judgment of the extent to which the sixth criterion is met is shown by the following corrected percentages of efficiency:

Information	100
Disarranged Sentences	82
Reading	90
Completions	87
Word Knowledge	98

Considering now the entire scale in this way we may compute the ratio of the total opportunities for skipping to the total amount of skipping actually done. This is 89. That is, 89 per cent of the possible gaps are filled. (See Chart I.) If there were no gaps the scale would be perfect, and the percentage would be 100.

The seventh criterion is not so easily met. If it were a perfect scale, then all those who cheated at least once should have done it on the last step, all those who cheated at least twice should have done it on the two easier steps, etc. The proportion actually doing this is shown at the bottom of the table, *viz.:*

PERCENTAGE OF THOSE CHEATING

At least once who cheat on step 1	96
At least twice who cheat on steps 1 and 2	81
At least three times who cheat on steps 1, 2 and 3	73
At least four times who cheat on steps 1, 2, 3, and 4	86
At least five times who cheat on steps 1, 2, 3, 4, and 5	100
Six times who cheat (necessarily) on steps 1, 2, 3, 4, 5 and 6	100

There remains, then, the eighth criterion involving the psychological significance of the steps. Does the scale actually represent a graded series of attitudes toward cheating? Specifically, does the fact that a person who will erase ink in order to cheat mean that he is more inclined toward cheating than one who will erase a sentence written in pencil, or does it mean that he is simply more stupid? Or again, is the person who cheats on the Reading test by erasing a phrase, more inclined toward cheating than the one who makes various changes in check marks on the Spelling test, or is he simply more interested in Reading? Fortunately, we have intelligence test scores on these children and we can satisfy one part of the criterion by reference to the correlation between IQ and scale score. But first it is necessary to translate the scale into a score in the case of each individual.

SCORING THE SCALE

Finding a satisfactory method of scoring performance scales of this sort has long been a source of annoyance to educational psycholo-

gists.[1] The primary difficulty involved is that of scoring those who skip one or more levels of the scale. If there were no skips, the score would be some value assigned to the highest level reached by the pupil tested. A way of arriving at such scale values is described above and the results are given in Table III.

It will be recalled that these values are multiples of SD which locate the position on the base line of a normal curve beyond which lie the particular percentages of cases found cheating at the levels represented by the six tests, and calling −3 SD, or the extreme anti-cheating attitude, zero. Theoretically it would be possible to fill in as many levels as one chose between those now occupied by the six tests. Assuming that this has been done, and that there is a level at each 0.1 SD, the scale would run as follows, with the present levels shown in italics:

1 2 3.....16 17 18 19 *20* 21 22 23 *24* 25 26 *27*

28 29 30 *31* 32.....39 *40* 41.....49 50 *51*

On such a completed scale, the score should be either the height reached, or a computed height in case there were several tasks at each level of difficulty, in which case the level reached might be regarded as that on which 50 per cent of the tasks were accomplished.

A modified score, allowing for imperfections in the scale, would be the total number of tasks completed, calling each task a unit. Adopting this plan in the case of the scale under discussion we have 20 as the score of one who reaches level 20, and 31 as the score of the one who reaches level 31, etc., assuming that in each case he has done the hypothetical tasks between. Thus one who misses the top level, makes only 40, or 11 less than the one who reaches the top.

If the scale were filled in, then to miss step 24 would be no more significant than to miss step 25, and the score of one who skipped the second step but reached the third would be the sum of all the tasks thus accomplished up to 31, that is, 30. But the scale is not filled in, and so the assumption is arbitrarily made that one who skips a step misses also all the hypothetical tasks between that step and the next lower one.

Thus, again, one who skips the top step, makes only 40, or 51 − 11, and one who skips the two top steps makes only 31, or 51 − 11 − 9. Likewise one who reaches the top step but skips the fifth, makes them all up to 31, but misses those from 32 to 40 inclusive, and then makes

[1] *Cf.* McCall: "How to Measure in Education." Pp. 300–306.

all from 41 to 51. His score is therefore 31 + 11 (40 to 51) or 51 – 9 (31 to 40) or 42. This is graphically represented as follows:

1	2	3
Arithmetic-spelling	Word Knowledge	Completions
.....18 19 **20**	21 22 23 **24**	25 26 **27**
4	**5**	**6**
Reading	Sentences	Information
28 29 30 **31**	32 33.....39 **40**	41 42.....50 **51**

Skipping step two, or three or four is thus seen to count less in the score than skipping step five or six, as there are fewer hypothetical tasks omitted.

Practically, the score is found by adding up the *intervals* that *precede* the *levels* actually reached. It is convenient, therefore, to simplify these intervals. This can be done without seriously changing their relative size by making the lowest level 21, the fourth 30, the fifth 39, and the sixth 52. They thus become multiples of three and can be divided by three to make, one, one, one, three, four.

The interval below 20 has still to be scored. At present this is the entire distance from −3 SD or from the extreme *anti* position. This gives too much weight to the first interval as compared with the rest. Adopting the suggestion made above in the discussion of the fourth criterion, and calling the neutral point zero, and assuming further that this is about as far below the Arithmetic-spelling level as this is below the Word Knowledge level, we have, say, −1.2 SD as the zero point, which equals 18 half SD's from −3. SD, which is 3 below 21, the present Arithmetic-spelling level. This step divided by 3 thus has equal score value with the next three.

The previous scale values ,with neutral at 18, now read:

18—21—24—27—30—39—52

Shifting the zero point to 18, and inserting the intervals just now determined (the differences divided by 3) we have these new scale values:

0—1—2—3—4—7—11

Assuming again that the gaps in this scale might be filled in, the entire range represented by the present six tests is:

0—*1*—*2*—*3*—*4*—5—6—*7*—8—9—10—*11*

Attaching these scale positions to the tests and adding the score values just described, we have:

	Level	Score
Neutral Point	0	0
Arithmetic-spelling	1	1
Word Knowledge	2	1
Completions	3	1
Reading	4	1
No test	5	1
No test	6	1
Disarranged Sentences	7	1 — 3
No test	8	1
No test	9	1
No test	10	1
Information	11	1 — 4
Highest possible score		11

Thus an individual's score is zero if he did not cheat at all, and if he did cheat, it is the sum of the score values of the tests on which he actually made changes.

Relation of Intelligence to the Scale Score

We are now ready to tackle the question of whether the difficulty levels are functions of cheating attitude or of intellectual ability, which was raised by the eighth criterion.

The correlation between the scale scores and Stanford IQ's is −.40 for the 263 cases of dependent children. These IQ's range from 60 to 140 with a mean of 97. This corresponds to the correlation secured by extensive testing a year ago with unselected groups and constitutes a rough validation of the scale.

Four interpretations suggest themselves concerning possible relations between intelligence and the scale.

First, that the increasing difficulties are *intellectual* in nature, thus requiring a higher intelligence for cheating on the high levels. Not only do the difficulties not appear to be of this character, involving only simple mechanical processes, but if this were the case the correlation would have been either zero or slightly positive since we already know, from previous testing, that the correlation between intelligence and unscaled cheating is between −.40 and −.50. That is, the tendency for the lower intelligence levels to cheat more would have been neutralized on the scale by the increasing difficulty of cheating. But this effect is evidently not present.

Second, that the more intelligent may see through the procedure or be suspicious enough not to cheat. While it is probable that some such factor is at work, the evidence lends no support to the theory, for all those who erased ink have IQ's of over 100, and the child with the highest IQ in the group, 140, cheated on all tests except the ink-erasing one.

Third, that the more intelligent do not need to cheat and so are less tempted to do so. This is plausible, as is also the fourth interpretation, that there is an inherent relationship between the traits involved. That is, it may be the case that lower intelligence is accompanied by either less resistance to cheating tendencies or by either inherited or acquired drives toward cheating, or, perhaps, toward the sort of adjustment of which cheating is an instance.

It should be noted, however, that the standard error of prediction is .9 of the SD of the distribution, which would make any prediction concerning the step on the scale of six steps that would be reached by a given IQ wholly unreliable.

The following details bear out the above discussion. The median IQ's of those reaching the various levels are:

Information	107
Disarranged Sentences	91
Reading	93
Language Completions	96
Word Knowledge	96
Arithmetic-spelling	105
Not cheating at all	103

Thus there is a slight tendency for the median IQ to increase as we descend the scale. These same facts are shown in a different way by computing the percentage of cases at each scale level who have IQ's of over 97, the mean of the whole group. They are:

Information	100
Disarranged Sentences	32
Reading	50
Language Completions	42
Word Knowledge	45
Arithmetic-spelling	73
Not cheating at all	62

We may conclude, therefore, that the steps on the scale are not particularly influenced by the differences in intelligence.

It might be said that these levels, instead of representing genuine attitudes toward cheating, merely represent degrees of recklessness or

daring on the part of the subject. Thus the person who erases the ink may have the same attitude toward cheating as the one who only makes a few changes in the spelling test, but is much more daring. Thus the scale tests "nerve" or willingness to take a chance rather than attitude toward cheating.

There are three ways of meeting this criticism. One is by an empirical check up with a series of paper and pencil tests of attitudes. The second is theoretical. We may argue that the pupil who is willing to take the chance is more set toward cheating than the one who will take only safe risks. This amounts to saying that for *practical purposes* this element of nerve or daring and attitude toward cheating may be regarded as one and the same thing.

The third check is by the use of other types of test material in which the mechanics are such that cheating is accomplished by the same general procedure, thus eliminating differences due to variation in the risk. This has been done and these other tests also fall into the same sort of scale, thus eliminating the theory of nerve.

Finally, do other attitudes enter in? Probably they do to a certain extent. Some tests are more interesting than others. Some are more closely related to school subjects and therefore have certain acquired drives not available for other material. This might be true more particularly of the arithmetic test as compared with the language completion. But here, again, for practical purposes this does not matter save as these attitudes vary among the different pupils. As all have been subjected to the same school discipline, there is not likely to be much difference. Furthermore, it should be noted that the acquired drives would not function in terms of cheating unless there was also the tendency to cheat. They operate merely as unknown motives, therefore, and can be practically ignored so far as the scale is concerned, although they are of course theoretically important in the analysis of the individual.

It remains now to emphasize the main thesis of this paper, namely, that the concept of a scale is general and may be applied to all kinds of testing. Suppose this scale for testing cheating in classroom situations be perfected. Suppose further that another be constructed for games or play situations; then another for situations involving money or business transactions. These could be placed alongside each other so that at each level there would be two or more situations. Such a scale would give a measure of the whole complex of behavior tendencies called dishonesty. The same procedure could be followed in

constructing scale, for other tendencies. Both positive as well as negative trends would be included. Situations involving all kinds of attitudes would be selected and graded. The outcome would be a general scale with many symptomatic situations at each level. A total character score or index might be obtained from the general level reached on the scale, or a more detailed picture of the relation of these trends to one another and of character as a whole might be revealed in the profile or some statistical coefficient.

Unless we succeed in getting down to basic dynamic factors underlying attitudes, it is doubtful if such scales as are here proposed can ever be as universal as the intelligence scales. They will forever rest on the performance of the groups from which they are made. But at the very least, they would serve as useful instruments for the quantitative description of individual behavior in relation to the behavior of stated groups.

A PERSONAL CONSTANT

H. HEINIS

University of Geneva, Geneva, Switzerland

Every measure of intelligence claims to answer the question: *Is it possible to recognize, even during the years of development, the degree of intellectual capacity an individual will attain to when once his evolution is ended?* From a more general point of view the question may thus be stated: *Is it possible at any age to determine all the intellectual potentiality of an individual?* Such is the problem that the present article will try to solve.

It is interesting to examine how the study of facts has brought one, little by little, nearer the solution. A first attempt was made by Binet when he introduced a precise formula into psychology: the *difference* between the actual age and the mental age, to define backwardness. This was, as we now realize, a very rough approximation, but it was one and it pointed the way to the true solution. The defects of this first attempt were soon recognized; and the fact, for which we are indebted to two psychologists, F. Kuhlmann and W. Stern, of taking instead of this difference, the *quotient* of these same ages, marks a real progress in the solution of the problem. A proof of it is that the variations of these quotients, concealed by the inexactitude of the experimental results, could not be determined for some time. W. Stern even came to the conclusion, that his new method of interpretation of intelligence examinations represents a personal constant. The criticisms stated by Bobertag[1] as regards this hypothesis were certainly justified, but the proofs advanced being insufficient, they in no way helped to reach the solution of the problem. Facts were needed, and a further ten years of experimentation were necessary before decisive results could be gathered together. The most favorable ground for this kind of research is afforded by the evolution of the abnormal, because the rate of their mental development is slower. They enable us to procure far more precise measures than those afforded by our means of investigation on the normal. By way of comparison one may say that the abnormal are for the study of physical evolution, what the inclined plane is for the study of the fall of bodies.

[1] Bobertag, O.: Uber Intelligenzprüfungen. *Zeitschr. f. angew, Psychologie*, 1912, p. 536.

Many well-known experimenters have tried to solve this problem by re-examining the same subjects at intervals, sufficiently great to avoid the influence of the preceding examinations. In most of these researches, however, the number of subjects examined was too limited. There is one exception: F. Kuhlmann and his assistants of the Minnesota-School for Feeble-minded examined 639 subjects from 1 to 20 years of age.[1] Amongst these subjects were represented all degrees of intelligence from idiocy to the normal. The examinations were repeated every two years by methods which were twice completely revised, first in 1912 and then in 1917.

The author's aim was to study the intellectual quotient. He came to the following conclusions:

1. *The mental ages of the feeble-minded increase with age at a rate proportionate to the degree of mental deficiency.*

2. *On the whole, the mental age ceases to increase between the ages of* 15 *and* 16, *the idiot grade ceasing development about* 3 *years earlier than the borderline grade.*

3. *The intelligence quotient decreases with age, and more for the higher than for the lower grades.*

4. *The rate of decline of the intelligence quotient increases some with age.*

Before being able to accept these assertions as established facts, it is necessary to be sure of the value of the scale employed for the measurements, which in this case is the metrical scale of Binet and Simon (we understand by that, all methods of intelligence investigation based on the principle of Binet's system). One has got into the habit of considering this method as a means of absolute measure, without its ever having been proved. It is true that Binet's genius anticipated the scientific knowledge of his time, thus making an immediate control of his work impossible. But now, as it is a question of establishing facts, this difficulty must be solved.

In the course of our researches for a method by which it may be possible to examine the intelligence during puberty, we have had to analyze the metrical scale of Binet and Simon.[2] Our study has shown us that this method, because of its rectilineal construction, is very

[1] Kuhlmann, F.: The Results of Repeated Mental Re-examinations of 639 Feeble-minded over a Period of Ten Years. *Journal of Applied Psychology*, Vol. V, pp. 195–224.

[2] Heinis, H.: La loi du developpement mental. *Archives de psych.*, No. 74, 1924, pp. 97–128.

limited in its applications; it represents a good approximation between 6 and 10 years of age, it is sufficiently exact for the ages from 3 to 12 to give valid results, but it cannot be used for the ages after puberty.

Considered from this point of view the assertions made by Kuhlmann correspond to the reality only insofar as they relate to childhood. The supposed cessation of physical growth between 13 and 16 years of age is thus due to the fact that the method employed is incapable of registering the small progress made during these years of development.[1]

Naturally such false appearances must have hidden from the author the true solution of the problem, without however hindering him from seeing the line he ought to follow to reach his aim.

Finally, F. Kuhlmann comes to this conclusion: that "*the several traits of the intelligence quotient are all due chiefly or possibly entirely to a decrease with age in the yearly increment in mental growth of average normal children.*" This has allowed the author to find a more adequate form—of course within the above mentioned limits—by which to express the variations of this quotient. He has tried to determine indirectly the median growth curve based on this result of his investigation: that for the different degrees of intelligence the intellectual quotient undergoes a diminution remaining (almost) constant from year to year. The "general type of median growth curve" would finally be expressed by the following formula:

$$y = A + (A - a) + (A - 2a) + \ldots\ldots\ldots\ldots\ldots (A - xa);^{2}$$

[1] The same criticism evidently applies also to all measurement of intellectual age taken on adolescent or adult subjects, with methods constructed on the Binet principle. No matter how much these measurements may agree together, they cannot procure any confirmation, for it is like turning in a circle to employ the same method to control the result. If the determination of the average mental age of the American army recruits (13.8 years) is contrary to common sense, there is nothing in that to surprise us. What is more astonishing is that this assumption has not provoked more protestation on the part of psychologists. Considering the fact, that the brain continues its growth after 16 till 30 years of age and even later, Porteus draws the following conclusions: "*If mental growth ceases at fifteen years of age, then this additional braingrowth has no connection with the growth of intelligence. It is a most unlikely assumption. Maybe the psychologist is like the man who went fishing and because he could not catch a fish came home with the firm belief that there were none to catch.*" Porteus, S.: Temperament and Mentality in Maturity, Sex and Race. *Journal of Applied Psychology*, Vol. VII, March, 1924, pp. 59–60.

[2] Kuhlmann only explains the construction of this curve. We express it in algebraic language.

where x = age and A = degree of development attained by the end of the first year after birth and $a = \frac{A}{16}$. Therefore, normal growth would be accomplished according to an arithmetical progression. At the age of 16 the maximum would be reached. To obtain the curve of growth for each degree of intelligence it would be sufficient to multiply by this degree.[1]

These are empirical rules which correspond very exactly to the facts observed, if one keeps within the above mentioned limits, but which are not satisfactory because they are too artificial.

By a quite different method and without any knowledge of these researches, we have been brought to conclusions little different from Kuhlmann's. What allowed us to anticipate them is a method of examining intelligence entirely different to that of Binet and his followers. By its means we are able to establish directly the curve of normal growth. Note that this was impossible by the Binet system, precisely because it is based on a principle that ignores the real psychological distances between the different ages. This may be realized by use of unities of unequal value. The alternative tests which form the metrical scale of Binet-Simon are unities of this kind.

We have fixed as our aim the determination of the quantitative law of the normal development of human intelligence. The metrical scale being of no use for similar investigations, it was necessary to construct a new one and to base this new method on real unities. Such unities are found in the graduated tests; so one unique series of rationally graduated tests, applicable to all ages would satisfy our ambitions provided that each test measures intelligence and that all together they should represent a sufficiently complete means of intellectual investigation. For our study we have been able to use the results obtained by A. Descoeudres[2] for the ages of 2 to 7 years, those of Vermeylen[3] for the ages of 6 to 12 years and our own (unpublished) researches on adolescents for the ages of 13 to 18 years.

These three investigations taken *in globo* include all ages from 2 to 18 years; that is they embrace almost all the years of mental develop-

---

[1] Here by degree of intelligence is meant the intelligence quotient of a subject examined at one year of age.

[2] Descoeudres, A.: "Le développement de l'enfant de 2 à 7 ans." Neuchâtel et Paris, 1921.

[3] Vermeylen, Dr. G.: Les débiles mentaux. *Bulletin de l'Institut général de psychologie*, 1922.

ment. In consequence one could expect to obtain a solution of the problem by the synthesis of these experimental results. In fact, the values V found by experiment will be satisfying to the equation:

$$V = a - be^{\frac{x}{d}},$$

where x = age, a, b, d, represent positive constants and e indicates the base of the natural logarithm.

From this it results that the law of mental development will be expressed by the logarithmic:

$$Y = B(1 - e^{\frac{x}{d}}) \qquad (1)$$

To give this expression a more concrete form we may replace the letters by the values calculated from the results given by Vermeylen's system and we obtain (see figure):

$$Y = 429(1 - e^{\frac{x}{6.675}}) \qquad (2)$$

According to these conclusions, one can see that the growth of human intelligence does not normally stop from the time of birth till old age. This progress though remaining positive diminishes from year to year, so that by the age of 40 the normal man has practically attained his maturity. As to the child, his development is rapid especially during the first years. Already at the age of 5 he has accomplished half his mental evolution. It is thanks to the mathematical interpretation, to which they lend themselves, that these indications may be expressed in such a precise way. Here is another example of this interpretation:

P. Ballard,[1] having tried in vain to measure yearly progress after 16 years, has proposed to space the ages and to test only representatives of 20, 30, and 40 years of age. Here is the question we would ask: *Could a method based on Binet's principle serve in this case as a means of measurement?*

The following answer is given by formula (2).

Between 16 and 20 years the progress of development is 18 degrees
Between 20 and 30 years the progress of development is 15 degrees
Between 30 and 40 years the progress of development is 3 degrees

[1] Ballard, P. B.: The Limit of the Growth of Intelligence. *British Journal of Psychology*, October, 1921, p. 138. It is just possible that the growth of intelligence after 16, imperceptible as it is from year to year, may in 10 years reach an amount that may be easily measured.

On the other hand the minimal psychological distance that can be measured by the metrical scale of Binet-Simon, is that of the development from 9 to 10 years. It is equal to 15 degrees. From this we conclude that it is perfectly possible, even with Binet's system, to measure how much in advance the subjects of 20 are with regard to those of 16 and even those of 30 years in comparison with the first. On the contrary, the small increase acquired from 30 to 40 escapes the precision of the above mentioned method of measurement.

To complete this discussion it may be useful to insist somewhat on the point of origin of our curve. The formula gives for

$$x = 0, \quad y = 0$$

which means, that at the time of birth the development of the intelligence is nil. This is a purely theoretical assertion, for one can hardly admit that the intelligence suddenly arises after 9 months of our existence. Are the senses of touch and taste, which appear already developed at birth, without influence on our mind?

We are totally ignorant, it is true, of what happens in our psychic nature from the moment of the fecundation of the egg till the awakening of the intelligence; but there must certainly be an evolution. It is therefore necessary to admit that during this preparatory phase, a course of development takes place, such, for instance as the curve indicated by the dots on our graph, more compatible with the principle of evolution than would be an abrupt origin. Whatever this curve may be, we have been able to show that formula (1) includes all the degrees of normal psychical growth. It follows that on our diagram the point of origin represents the absolute zero of development, the degree y being therefore absolute also. It is unnecessary to dwell upon the importance of this result. We shall only point out that this transformation of relative unities of measurement into absolute unities facilitates the solution of certain problems. In very many cases this transformation is indispensable. To make this point clear we may reproduce merely an example already given in the above mentioned study:[1]

Dr. Vermeylen has studied the relation existing in children between the activity of elaboration and of acquisition. The graphical representation of the results which he obtained gives two curves which intersect at a point corresponding to eight years of age. The author comes to the conclusion that: *Between seven and nine years of age there*

[1] *Loc. cit.*, pp. 124–127.

comes in the psychological evolution of the child a phenomenon sufficiently important to reverse completely the relation between the two great modes of mental activity. If one expresses the same results in absolute degrees, the respective positions of the two representative curves change totally and the phenomenon in question reduce itself to the fact that *in the development of the child the activity of elaboration holds a place of extreme importance.*

In exposing the conclusions to which F. Kuhlmann has been brought, we called attention to the fact, that the problem of the study of the IQ's variations is in fact the same as that of the study of the average growth curve. This curve being now established, the following thesis may be formulated:

All the properties of the IQ are deduced from the law of mental development.

This we shall show.

I. *Proposition.—The mental age of the feeble-minded increases with age at a rate proportionate to the degree of mental deficiency.*

To reach this conclusion, it is not necessary to pass by the IQ. If

$$y = b(1 - e^{-\frac{x}{d}})$$

indicates the average growth curve, that of an individual of any degree of intelligence—as we have been able to prove—may be deduced from it and takes the form

$$Y = B\,(1 - e^{-\frac{x}{d}}).$$

For the degrees superior to the normal $B > b$ and in consequence for any one of the inferior degrees, $0 < B < b$. The progress realized between two ages x_1 and x_2 is expressed by the difference:

$$Y_2 - Y_1 = B(1 - e^{-\frac{x_2}{d}}) - B(1 - e^{-\frac{x_1}{d}}$$

$$= B(e^{-\frac{x_1}{d}} - e^{\frac{x_2}{d}}).$$

The formula (2) gives d as superior to 1. Then

$$e^{-\frac{x_1}{d}} > e^{-\frac{x_1}{d}},$$

from which we obtain the inequality

$$B(e^{\frac{x_1}{d}} - e^{\frac{x_2}{d}}) > 0.$$

The bigger B, the more the left product will increase; *i.e.*, speaking generally, the mental age increases with age at a rate proportionate to the degree of intelligence.

II. *Proposition.—For the degrees lower than the average the IQ decreases with age, and more for the higher than for the lower grades.*

Let x be any given age, $(x + 1)$ the following year, X_x the intellectual age of an individual of x years; the value $\frac{X_x}{x}$ = his IQ. The first part of the above mentioned proposition will therefore give in algebraical language

$$\frac{X_{x+1}}{x+1} - \frac{X_x}{x} < 0.$$

How may we express this as regards the curve of normal growth? The abscissa X_x determines on this curve the ordinate

$$y = b(1 - e^{-\frac{X_x}{d}}).$$

This same value y must again come out in the curve of normal growth

$$y = B(1 - e^{-\frac{x}{d}}).$$

The first member of these two equations being equal, the second will also be so. Then

$$b(1 - e^{\frac{X_x}{d}}) = B(1 - e^{-\frac{x}{d}}). \qquad (3)$$

This equality allows us to express X_x in functions of x. Let us again put $\frac{B}{b} = c$ and we shall have after a few transformations

$$\frac{X_x}{x} = -D\frac{\log(1 - c + ce^{-\frac{x}{d}}}{x} \qquad (D = \text{constant})$$

Now let us formulate the difference between the intelligence quotient corresponding to two years chronologically consecutive:

$$\frac{X_{x+1}}{x_{+1}} - \frac{X_x}{x} = -D\frac{\log(1 - c + ce^{-\frac{x+1}{d}})}{x+1} + \frac{log(1 - c + ce^{-\frac{x}{d}})}{x}$$

The second part of this equality may be transformed by this product:

$$= \frac{D}{x(x+1)} \log \frac{(1 - c + ce^{-\frac{x}{d}})^{x+1}}{(1 - c + ce^{-\frac{x+1}{d}})^{x}}.$$

According to the indications of formula (2) $d > 1$, then

$$0 < e^{-\frac{1}{d}} < 1.$$

It follows in the same way

$$0 < (e^{-\frac{1}{d}})^{x} < 1.$$

As the c in the present case confines itself to the limits

$$0 < c < 1;$$

the value $ce^{-\frac{x}{d}}$ and *a fortiori* $ce^{-\frac{x+1}{d}}$ remain a negligible quantity as regards c. To discuss the sum of $1 - c + ce^{-\frac{x}{d}}$ or that of $1 - c + ce^{-\frac{x+1}{d}}$ it is sufficient to take into consideration the difference $(1 - c)$ which in its turn reduced the quotient. It is then necessary to take the logarithm (formula (3)) in this same expression $(1 - c)$.

For the degrees of intelligence in question this difference places itself between the limits

$$0 < 1 - c < 1.$$

In consequence its logarithm will always take a negative value, that is

$$\frac{X_{x+1}}{x+1} - \frac{X_x}{x} < 0.$$

The nearer the value $(1 - c)$ is to zero, the greater is the inequality or, which comes to the same thing, the greater the decrease of the intelligence quotient. This means that the intelligence quotient decreases more for the higher grades.

III. *Proposition.—The rate of decline of the intelligence quotient increases with age.*

We refer to what we have just seen with regard to formula (3), that the discussion of the numeral value may be reduced to that of the following expression

$$-D\frac{\log(1-c)}{x(x+1)}.$$

The rate of decrease between two consecutive years

$$\left(\frac{X_x}{x} - \frac{X_{x+1}}{x+1}\right) - \left(\frac{X_{x+1}}{x+1} - \frac{X_{x+2}}{x+2}\right)$$

will be established from the difference

$$D\frac{\log(1-c)}{(x+1)(x-2)} - D\frac{\log(1-c)}{x(x+1)}.$$

After simplification, this may be reduced to a formula easy to evaluate:

$$-\frac{2D\log(1-c)}{x(x+1)(x+2)}.$$

This expression will always give a positive value, as the logarithm $(1 - c) < 0$.

IV. *Proposition.—The intelligence quotient of cases above the average will increase with age instead of decreasing.*

This is said when

$$c > 1, \frac{X_{x+1}}{x+1} - \frac{X_x}{x} > 0.$$

In this case the previous reasoning may not be directly applied as $(1 - c) < 0$, and there is no logarithm of a negative numerical value. But in considering the symmetrical construction of the equality (3) it will be seen that by changing the variables x and X_x deductions similar to the preceding ones may be drawn. By expressing this time x in function of X_x, the following inequality is therefore reached:

$$\frac{x+1}{X_{+1}} - \frac{x}{X_x} < 0$$

Which supposes that

$$\frac{x+1}{X_{x+1}} < \frac{x}{X_x};$$

in consequence

$$\frac{X_{x+1}}{x+1} > \frac{X_x}{x}$$

and

$$\frac{X_{x+1}}{x+1} - \frac{X_x}{x} > 0.$$

The relation between the rate of growth and the degree of intelligence has therefore been reversed by the introduction of the new function $x = f(X_x)$. This fact and the preceding result bring us to the proposition:

V. *Proposition.—The intelligence quotient of cases above the average increases with age and more for the higher than for the lower grades.*

For the same reason we come to the conclusion:

VI. *Proposition.—The rate of growth of the intelligence quotient increases with age.*

These deductions, which are, as can be seen, in perfect agreement with Kuhlmann's experimental results, prove once more the truth of our law. What complicated the calculations, was precisely the notion of the IQ conforming so little with the logarithmical progress of mental development. It would certainly be better to do away with this interpretation of intelligence examinations and to try to establish something more adequate. This is easy, if one takes as starting point, not the metrical scale of Binet and Simon, but a method such as the one we have proposed: *i.e.*, a method based on a unique series of rationally graded tests.

As has already been said the curve of individual growth may be expressed in this case by the formula:

$$Y = B(1 - e^{-\frac{x}{d}}),$$

if that of the average growth be

$$y = b(1 - e^{-\frac{x}{d}}).$$

Let us take these equations member by member and we shall obtain:

$$\frac{Y}{y} = \frac{B}{b} = c.$$

It is this constant value c, invariable for any given individual, which characterizes permanently the relation existing between the development of the individual and the normal growth curve. We have called this value the *personal coefficient.*

Definition.—The personal coefficient of any given individual is equal to the result of the intelligence examination divided by the normal degree of intelligence corresponding to his age, both measures being given in absolute graduation.

We shall now try to verify the constancy of the new interpretation of intelligence examination by Kuhlmann's experiments.

These results have been obtained, as we have already seen, by a metrical scale. They are therefore given in real and mental ages. In consequence, if they are to be used for our ends, a connection must be established between Binet's system with ours.[1]

---

[1] To facilitate the use of the point scale, Yerkes and Wood had proposed a method of interpretation analogous to that of the IQ. To distinguish it from the latter, they had called it *coefficient of intelligence.* This CI is the quotient of the sum of points obtained by their method with a subject by that which one would normally expect to obtain with the same individual. The practical results of this proposition have been shown to be inferior to those of the intelligence quotient. As we can today realize, this inferiority is due to the fact that the number of these points was established in too arbitrary a fashion. The IQ is calculated on more absolute values, *i.e.*, the ages.

Between these two methods of interpretation, the IQ and the CI, can be placed the *intelligence ratio* (IR) of the Myers mental measure. Its numerical value is obtained by dividing the raw score by the age-in-months. The IR certainly marks a progress on the CI. A more perfect method of interpretation of the Myers mental measure ought, however, to be found. Let us take the years from 7 to 14 for which the said method appears to be, except for some approximations, rationally graded. Following the formulas indicated in our work (*loc. cit.*, p. 125), let us

Let us suppose X_2 = intellectual age, given to a case of x years old by Binet-Simon's metrical scale. All other things being equal, we reach by our method the following result:

$$Y = b(1 - e^{-\frac{x_2}{d}}).$$

The number of degrees corresponding normally to the age of the subject will be:

$$y = b(1 - e^{-\frac{x}{d}})$$

and the quotient

$$\frac{Y}{y} = \frac{1 - e^{-\frac{x_2}{d}}}{1 - e^{-\frac{x}{d}}} = c$$

represents the personal constant. Replace d by its numeral value, which is 6675 (equality 2):,

$$c = \frac{1 - e^{-\frac{x_2}{6.675}}}{1 - e^{-\frac{x}{d}}}$$

We have here the definite formula which allows us to calculate the PC of any subject whose intellectual age has been determined by any given method.

Example.—An abnormal child of 13 years 8 months ($x = 13.67$) has been found to have the intelligence of a child of 5 years 3 months ($X_2 = 5.25$). *The personal coefficient* will be:

$$c = \frac{1 - e^{-\frac{5.25}{6.675}}}{1 - e^{-\frac{13.675}{6.675}}} = 0.626.$$

Before applying this formula to Kuhlmann's experimental results, it is necessary to examine them closely. They must be judged according to the way by which they have been obtained.

The author verified twice his metric scale on 1000 school-children. The last revision aimed at determining the errors due to the general

---

change the relative graduation g to an absolute graduation G. This may be done by the formula:

$$G = g + 91$$

Let g_m be the raw score obtained by a subject of x years and g_x his normal raw score. The expression:

$$C = \frac{g_m + 91}{g_x + 91}$$

indicates the value of the *personal coefficient.*

tendency of the scale to measure too high or too low at the different levels. These errors once determined, the necessary corrections were made for all measures taken before the revision except for the ages under 4. Furthermore, considering the fact that Binet-Simon method is insufficient to measure cases over 15 years of age, we are therefore obliged to consider only the results obtained on the ages of 4 to 15. But within these limits we must give Kuhlmann's results the credit of really representing facts, perhaps better than many other investigations of this kind. There remains nonetheless one more cause of error which must be discussed, although it is of a secondary importance. It is due to a difficulty of classification.

As we have already indicated, Kuhlmann wished, by his investigation, to determine the successive values of the IQ of an individual in the course of his evolution. The best method would be evidently to follow the complete mental evolution of individuals, representing the different levels of intelligence. The measurements should be taken at regular and frequent intervals. Without doubt many factors hinder the execution of so vast a program. But for theoretical reasons the principle must be maintained. Let us see how the author applied it to the practical circumstances he had to deal with.

Because of change of method, the results of the examinations that could be compared with each other only covered a maximal period of five years. As these examinations were taken at intervals of two years, none of the cases furnished more than three IQ's. It was with these three values (the cases who offered less than three being eliminated) that the synthetical type of each intellectual level was to be determined. According to custom, Kuhlmann distinguished four intellectual levels and defined them for the lowest age (7) where these levels differ sufficiently, in the following way:

1. (Idiots) the IQ amounts to 0 — 0.24
2. (Imbeciles) the IQ amounts to 0.25 — 0.49
3. (Morons) the IQ amounts to 0.50 — 0.74
4. (Borderline) the IQ amounts to 0.75 —

All cases 7 years old at the first examination were classed according to these definitions. The average of the IQ (AIQ) found for the representatives of each group, allowed characterization of the four synthetical types for the ages of 7, 9 and 11. By simple interpolation the AIQ for 8 could be determined and served as a basis for the classi-

fication of the cases who were 8 years old, when they were examined for the first time. Only one had to proceed by an inverse method: for this time the averages were given and the values to be determined were the limits containing the IQ's which furnished these averages. The 8-year-old cases were then added to the 7-year-old and in this way the chain could be continued.

One cannot contest the value of this method, proceeding by enveloping tangents in order to realize the principle above indicated. All the same, the precision of the values established in this fashion depends on the number of IQ's with which one calculated the averages. If this number is small, as was forcibly the case for the group (age 7) which was at the basis of the experiment, the results are too much influenced by individual diagnostics. Owing to the small amount of precision of the method employed, the results obtained are not free from notable errors. An example of this is seen in the case of the morons, where a group of backward rather than subnormal cases influenced unfavorably the determination of the IQ's of this degree. The best ascertained trait at the present time of the IQ is its diminution in the inferior degrees. But the IQ's of the group of morons aged from 7 to 8 increase. There must therefore be some irregularities in the PC of the moron category.

It is the same for the first two groups representing the idiots. But as they do not reach the mental age of 4, we are obliged to leave aside the results obtained for this degree of intelligence, and for the ages of 7 to 10 in the group of the imbeciles.

One difficulty remains to be stated. To translate IQ into PC, especially for the years of rapid development, exact knowledge of the chronological age of the subjects is necessary. The author gives us very summary information on this point. For instance, all individuals aged from 7 to 8 are classed as if they were 7 years old. Nevertheless, if the number of subjects is large enough, this does not matter so much, as the average will not vary much from 7.5 years. In the case of small groups, as with the borderline cases of 7 and 8 (6 and 9 cases) or with the morons of the same ages (11 and 10 cases) no age average can be determined with precision. For lack of necessary information we must forego the translation into PC's of the IQ's corresponding to the ages of 7 and 8. Taking again into consideration the upper limits that we have been obliged to fix to the method employed, we shall be able to draw from Kuhlmann's investigation those results that are really of use to us. They are confined to the

phase of development from 9 to 15. Table I gives the PC's calculated by the average mental and chronological age.

TABLE I

Borderline

Age	9.5	10.5	11.5	12.5	13.5	14.5	15.5
Number of cases	25	35	40	39	42	42	37
Average mental age	7.19	7.77	8.37	9.07	9.38	9.82	10.13
Personal coefficient	.871	.870	.872	.875	.871	.872	.874

Morons

Age	9.5	10.5	11.5	12.5	13.5	14.5	15.5
Number of cases	52	76	101	111	118	111	104
Average mental age	5.70	6.26	6.81	7.22	7.57	8.01	8.49
Personal coefficient	.758	.779	.780	.782	.783	.788	.797

Imbeciles

Age	9.5	10.5	11.5	12.5	13.5	14.5	15.5
Number of cases	...	59	83	93	104	99	99
Average mental age	...	3.85	4.07	4.32	4.46	4.64	4.75
Personal coefficient	...	.558	.560	.564	.562	.564	.564

In this account we have given three decimals for the PC. This is, indeed, too great a precision; for even if we leave on one side the probable error of the average mental age, the calculation of the PC had to be based on a more or less exact age average. This fact is enough by itself to render doubtful the exactitude of the third decimal. If one wishes to keep to really proved results, one cannot take into account in the indication of PC values any amount further than the second decimal. Table I is simplified thereby:

TABLE II.—PERSONAL COEFFICIENTS

Age	9.5	10.5	11.5	12.5	13.5	14.5	15.5
Borderline	.87	.87	.87	.87	.87	.87	.87
Morons	.76?	.78	.78	.78	.78	.78?	.79?
Imbeciles	.....	.56	.56	.56	.56	.56	.56

When we remind ourselves of what has been said as to the composition of the moron group, we notice that this table shows us the PC as

being in fact a real personal constant. Many other experimental results, especially the documents W. Stern has gathered in evidence of the constancy of his intelligence quotient support this assumption.[1]

From all this we may conclude that the personal coefficient has a permanent value. It therefore affords a most rational means of classifying human intelligences. We have attempted such a classification based on the more or less empirical results of W. Stern and Terman.[2]

CLASSIFICATION

Designation	Personal coefficient
Idiot	0–0.40
Imbecile	0.41–0.68
Moron	0.69–0.86
Feeble	0.87–0.96
Normal	0.97–1.02
Strong	1.03–1.12
Superior	1.13–1.30

As to the practical application of the new method of interpretation of intelligence examinations, it can be much simplified by the help of a table or better still an appropriate graphical representation. For the present, we believe we had better give neither the one nor the other. Though the experimental results of many point scales are known, it is none the less true that not one of them is yet rationally graded.[3] Too much attention has been paid to the creation of new methods and not enough to the study of those already existing. This is perhaps the principle reason for which till now all attempts to determine the curve of mental growth have failed.[4]

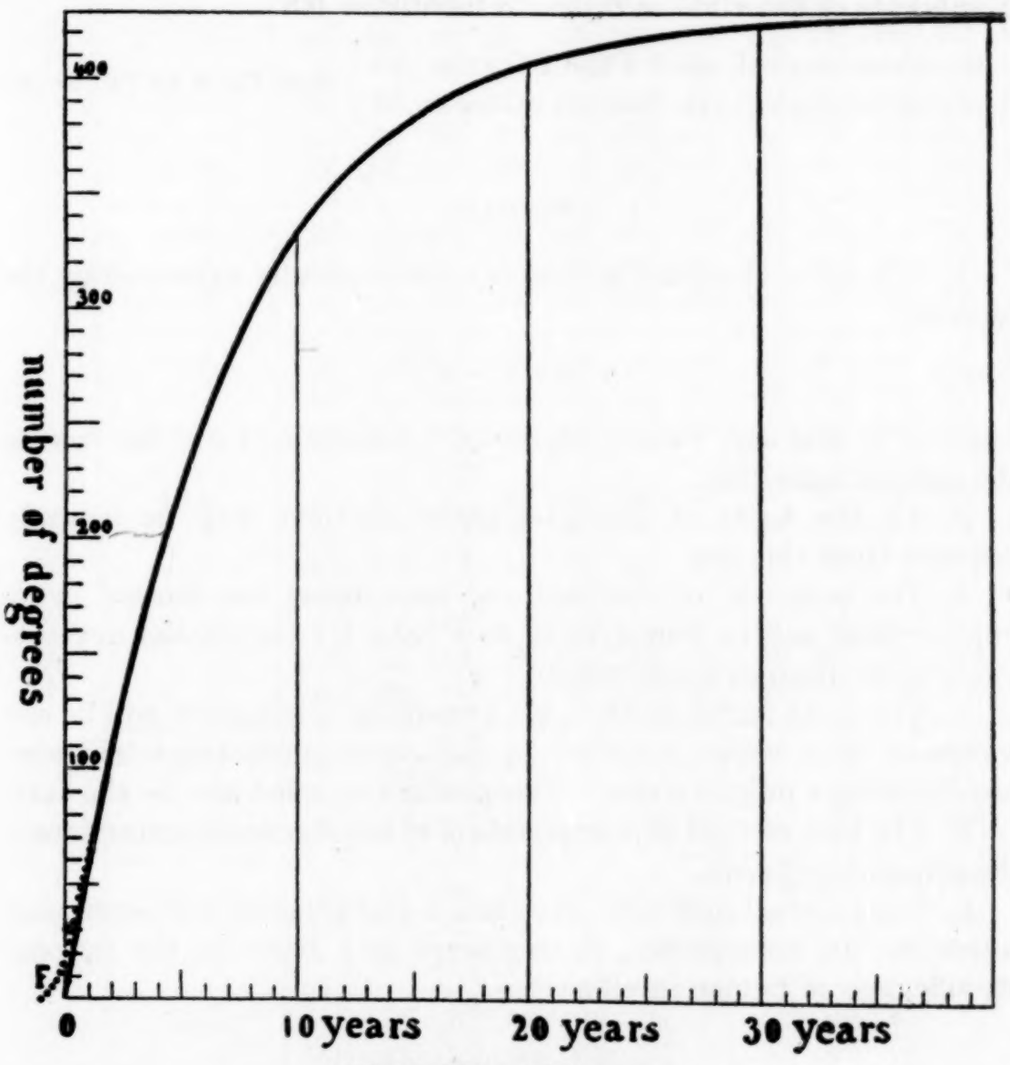
[1] *Loc. cit.*, p. 121.

[2] *Loc. cit.*, p. 123.

[3] G. Arthur and H. Woodrow have proposed an absolute intelligence scale based on absolute unities. The authors define the latter by the unities of the number resulting from the division of the yearly progress of the averages of a test, by the middle standard deviation. It is easy to show with formula (1) that such a graduation cannot be a rational one, as the scores it gives do not satisfy the logarithmic equation. Arthur G., and Woodrow, H.: An Absolute Intelligence Scale. *Journal of Applied Psychology*, June, 1919, pp. 118–137.

[4] In fact, F. Freeman after having discussed "the interpretation and application of the intelligence quotient" comes to the same conclusion (*cf. Journal of Educational Psychology*, Vol. XII, 1921, p. 13).

The metrical scale actually in use cannot be accused of this same defect. On the contrary, many of them are now the best studied instruments for the examination of intelligence. It therefore seems to us useful to render our new method of interpretation easily accessible to those who are accustomed rather to measuring the intelligence by a



Mental growth curve.
F = moment of the fecundation of the egg.

metrical scale. The formula indicated on page 174 evidently allows one to calculate the PC's, when the mental and real age of a subject are known. This calculation, it is true, is not very simple. A reference

table is therefore necessary. This is the reason why we have established that which follows.

The use of these instruments of calculation, the table "*à double entrée,*" is too well known for us to dwell upon them. Here is an example:

Example.—Chronological age of a child = 9 years 10 months = 9.10. Mental age of this child = 6 years 8 months = 6.8

$$\left\{\begin{array}{l}\text{the chronological age 8.9 the PC as } = .83\\ \text{the chronological age 10.0 the PC as } = .81\end{array}\right\} \text{thus for 9.10 PC} = .82$$

Summary

1. The curve of mental growth is mathematically expressed by the formula:

$$y = b(1 - e^{-\frac{x}{d}})$$

where x = real age, b and d represent constants and e = the base of the natural logarithm.

2. All the traits of the intelligence quotient may be logically deduced from this law.

3. The principle of the metrical scale being too limited in its improvement can no longer serve as a basis for the ulterior development of intelligence examination.

4. The most fertile method for measuring intelligence will be one composed of a happy selection of rationally graduated intelligence tests forming a unique series. This graduation must also be absolute.

5. The best method of interpretation of intelligence examinations is the *personal coefficient.*

6. The *personal coefficient* represents a real personal and permanent constant. In consequence, it may serve as a basis for the rational classification of human intelligence.

MENTAL AGE

Chronological Age		3						4					
Years / Months		0	2	4	6	8	10	0	2	4	6	8	10
	6–0	.61	.64	.66	.69	.71	.74	.76	.78	.80	.83	.85	.87
	4	.59	.62	.64	.67	.69	.71	.74	.76	.78	.80	.82	.84
	8	.58	.60	.62	.65	.67	.69	.71	.73	.76	.78	.80	.81
	7–0	.56	.59	.61	.63	.65	.67	.69	.71	.73	.75	.77	.79
	4	.55	.57	.59	.61	.63	.65	.68	.70	.72	.73	.75	.77
	8	.53	.55	.58	.60	.62	.64	.66	.68	.70	.72	.74	.75
	8–0	.52	.54	.56	.58	.60	.62	.64	.66	.68	.70	.72	.74
	4	.51	.53	.55	.57	.59	.61	.63	.65	.67	.69	.70	.72
	8	.50	.52	.54	.56	.58	.60	.62	.64	.66	.67	.69	.71
	9–0	.49	.51	.53	.55	.57	.59	.61	.63	.65	.66	.68	.69
	4	.48	.50	.52	.54	.56	.58	.60	.62	.63	.65	.67	.68
	8	.48	.50	.52	.53	.55	.57	.59	.61	.62	.64	.66	.67
	10–0	.47	.49	.51	.53	.54	.56	.58	.60	.61	.63	.65	.66
	4	.46	.48	.50	.52	.54	.56	.57	.59	.61	.62	.64	.65
	8	.46	.47	.49	.51	.53	.55	.57	.58	.60	.61	.63	.65
	11–0	.45	.47	.49	.51	.52	.54	.56	.58	.59	.61	.62	.64
	4	.45	.46	.48	.50	.52	.54	.55	.57	.59	.60	.62	.63
	8	.44	.46	.48	.49	.51	.53	.55	.56	.58	.59	.61	.62
	12–0	.44	.45	.47	.49	.51	.52	.54	.56	.57	.59	.60	.62
	4	.43	.45	.47	.48	.50	.52	.53	.55	.57	.58	.60	.61
	8	.43	.45	.46	.48	.50	.51	.53	.55	.56	.58	.59	.61
	13–0	.42	.44	.46	.48	.49	.51	.53	.54	.56	.57	.59	.60
	4	.42	.44	.46	.47	.49	.51	.52	.54	.55	.57	.58	.60
	8	.42	.44	.45	.47	.49	.50	.52	.53	.55	.56	.58	.59
	14–0	.42	.43	.45	.47	.48	.50	.51	.53	.54	.56	.57	.59
	4	.41	.43	.45	.46	.48	.50	.51	.53	.54	.56	.57	.58
	8	.41	.43	.44	.46	.48	.49	.51	.52	.54	.55	.57	.58
	15–0	.41	.43	.44	.46	.47	.49	.50	.52	.53	.55	.56	.58
	4	.41	.42	.44	.46	.47	.49	.50	.52	.53	.55	.56	.57
	8	.40	.42	.44	.45	.47	.48	.50	.51	.53	.54	.56	.57
	16–0	.40	.42	.43	.45	.46	.48	.49	.51	.53	.54	.55	.57

MENTAL AGE

Chronological Age													
Years		5						6					
	Months	0	2	4	6	8	10	0	2	4	6	8	10
	6–0	.89	.91	.93	.95	.96	.98	1	1.02	1.03	1.05	1.07	1.08
	4	.86	.88	.90	.92	.93	.95	.97	.98	1	1.02	1.03	1.05
	8	.83	.85	.87	.89	.90	.92	.94	.95	.97	.99	1	1.02
	7–0	.81	.83	.85	.86	.88	.90	.91	.93	.94	.96	.97	.99
	4	.79	.81	.82	.84	.86	.87	.89	.90	.92	.93	.95	.96
	8	.77	.79	.80	.82	.84	.85	.87		.90	.91	.92	.94
	8–0	.75	.77	.79	.80	.82	.83	.85	.86	.88	.89	.90	.92
	4	.74	.75	.77	.79	.80	.82	.83	.84	.86	.87	.89	.90
	8	.72	.74	.76	.77	.79	.80	.81	.83	.84	.86	.87	.88
	9–0	.71	.73	.74	.76	.77	.79	.80	.81	.83	.84	.85	.86
	4	.70	.71	.73	.74	.76	.77	.79	.80	.81	.83	.84	.85
	8	.69	.70	.72	.73	.75	.76	.77	.79	.80	.81	.83	.84
	10–0	.68	.69	.71	.72	.74	.75	.76	.78	.79	.80	.81	.83
	4	.67	.68	.70	.71	.73	.74	.75	.77	.78	.79	.80	.82
	8	.66	.68	.69	.70	.72	.73	.74	.76	.77	.78	.79	.80
	11–0	.65	.67	.68	.69	.71	.72	.73	.75	.76	.77	.78	.79
	4	.65	.66	.67	.69	.70	.71	.73	.74	.75	.76	.77	.78
	8	.64	.65	.67	.68	.69	.71	.72	.73	.74	.75	.77	.78
	12–0	.63	.65	.66	.67	.69	.70	.71	.72	.73	.75	.76	.77
	4	.63	.64	.65	.67	.68	.69	.70	.72	.73	.74	.75	.76
	8	.62	.63	.65	.66	.67	.68	.70	.71	.72	.73	.74	.75
	13–0	.61	.63	.64	.65	.67	.68	.69	.70	.71	.73	.74	.75
	4	.61	.62	.64	.65	.66	.67	.69	.70	.71	.72	.73	.74
	8	.61	.62	.63	.64	.66	.67	.68	.69	.70	.71	.73	.74
	14–0	.60	.61	.63	.64	.65	.66	.68	.69	.70	.71	.72	.73
	4	.60	.61	.62	.64	.65	.66	.67	.68	.69	.70	.72	.73
	8	.59	.61	.62	.63	.64	.65	.67	.68	.69	.70	.71	.72
	15–0	.59	.60	.62	.63	.64	.65	.66	.67	.68	.69	.71	.72
	4	.59	.60	.61	.62	.63	.65	.66	.67	.68	.69	.70	.71
	8	.58	.60	.61	.62	.63	.64	.65	.67	.68	.69	.70	.71
	16–0	.58	.59	.60	.62	.63	.64	.65	.66	.67	.68	.69	.70

MENTAL AGE

Chronological Age: Years	Months	7						8					
		0	2	4	6	8	10	0	2	4	6	8	10
	6–0	1.10	1.11	1.13	1.14	1.15	1.17	1.18	1.19	1.20	1.21	1.22	1.24
	4	1.06	1.08	1.09	1.10	1.12	1.13	1.14	1.15	1.16	1.18	1.19	1.20
	8	1.03	1.04	1.06	1.07	1.08	1.09	1.11	1.12	1.13	1.14	1.15	1.10
	7–0	1	1.01	1.03	1.04	1.05	1.06	1.08	1.09	1.10	1.11	1.12	1.13
	4	.97	.99	1	1.01	1.02	1.04	1.05	1.06	1.07	1.08	1.09	1.10
	8	.95	.96	.98	.99	1	1.01	1.02	1.03	1.04	1.05	1.06	1.07
	8–0	.93	.94	.95	.97	.98	.99	1	1.01	1.02	1.03	1.04	1.05
	4	.91	.92	.93	.95	.96	.97	.98	.99	1	1.01	1.02	1.03
	8	.89	.90	.92	.93	.94	.95	.96	.97	.98	.99	1	1.01
	9–0	.88	.89	.90	.91	.92	.93	.94	.95	.96	.87	.98	.99
	4	.86	.88	.89	.90	.91	.92	.93	.94	.95	.96	.97	.98
	8	.85	.86	.87	.88	.89	.90	.91	.92	.93	.94	.95	.96
	10–0	.84	.85	.86	.87	.88	.89	.90	.91	.92	.93	.94	.95
	4	.83	.84	.85	.86	.87	.88	.89	.90	.91	.92	.92	.93
	8	.81	.83	.84	.85	.86	.87	.88	.89	.90	.90	.91	.92
	11–0	.81	.82	.83	.84	.85	.86	.87	.88	.88	.89	.90	.91
	4	.80	.81	.82	.83	.84	.85	.86	.87	.87	.88	.89	.90
	8	.79	.80	.81	.82	.83	.84	.85	.86	.87	.87	.88	.89
	12–0	.78	.79	.80	.81	.82	.83	.84	.85	.86	.86	.87	.88
	4	.77	.78	.79	.80	.81	.82	.83	.84	.85	.85	.86	.87
	8	.76	.77	.78	.79	.80	.81	.82	.83	.84	.85	.86	.86
	13–0	.76	.77	.78	.79	.80	.81	.82	.82	.83	.84	.85	.86
	4	.75	.76	.77	.78	.79	.80	.81	.82	.83	.83	.84	.85
	8	.75	.76	.77	.78	.78	.79	.80	.81	.82	.83	.84	.84
	14–0	.74	.75	.76	.77	.78	.79	.80	.81	.81	.82	.83	.83
	4	.74	.75	.76	.76	.77	.78	.79	.80	.81	.82	.82	.84
	8	.73	.74	.75	.76	.77	.78	.79	.79	.80	.81	.82	.83
	15–0	.73	.74	.75	.76	.76	.77	.78	.79	.80	.81	.81	.82
	4	.72	.73	.74	.75	.76	.77	.78	.79	.79	.80	.81	.82
	8	.72	.73	.74	.75	.76	.77	.77	.78	.79	.80	.80	.81
	16–0	.71	.72	.73	.74	.75	.76	.77	.78	.78	.79	.80	.81

MENTAL AGE

Chronological Age		9						10					
Years	Months	0	2	4	6	8	10	0	2	4	6	8	10
	6–0	1.25	1.26	1.27	1.28	1.29	1.30	1.31	1.32	1.33	1.33	1.34	1.35
	4	1.21	1.22	1.23	1.24	1.25	1.26	1.26	1.27	1.28	1.29	1.30	1.31
	8	1.17	1.18	1.19	1.20	1.21	1.22	1.23	1.24	1.24	1.25	1.26	1.27
	7–0	1.14	1.15	1.16	1.17	1.18	1.18	1.19	1.20	1.21	1.22	1.23	1.23
	4	1.11	1.12	1.13	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20
	8	1.08	1.09	1.10	1.11	1.12	1.13	1.13	1.14	1.15	1.16	1.17	1.17
	8–0	1.06	1.07	1.08	1.08	1.09	1.10	1.11	1.12	1.12	1.13	1.14	1.15
	4	1.04	1.05	1.05	1.06	1.07	1.08	1.09	1.10	1.10	1.11	1.12	1.13
	8	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.07	1.08	1.09	1.10	1.10
	9–0	1	1.01	1.02	1.03	1.03	1.04	1.05	1.06	1.06	1.07	1.08	1.08
	4	.98	.99	1	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.06	1.07
	8	.97	.98	.99	.99	1	1.01	1.02	1.02	1.03	1.04	1.04	1.05
	10–0	.95	.96	.97	.98	.99	.99	1	1.01	1.01	1.02	1.03	1.03
	4	.94	.95	.96	.97	.97	.98	.99	.99	1	1.01	1.01	1.02
	8	.93	.94	.94	.95	.96	.97	.97	.98	.99	.99	1	1.01
	11–0	.92	.92	.93	.94	.95	.96	.96	.97	.98	.98	.99	.99
	4	.91	.91	.92	.93	.94	.94	.95	.96	.96	.97	.98	.98
	8	.90	.91	.91	.92	.93	.93	.94	.95	.95	.96	.97	.97
	12–0	.89	.90	.90	.91	.92	.92	.93	.94	.94	.95	.96	.96
	4	.88	.89	.89	.90	.91	.92	.92	.93	.94	.94	.95	.95
	8	.87	.88	.89	.89	.90	.91	.91	.92	.93	.93	.94	.94
	13–0	.86	.87	.88	.88	.89	.90	.91	.91	.92	.92	.93	.93
	4	.86	.86	.87	.88	.89	.89	.90	.90	.91	.92	.92	.93
	8	.85	.86	.86	.87	.88	.89	.90	.90	.91	.91	.92	.92
	14–0	.84	.85	.86	.87	.87	.88	.89	.89	.90	.90	.91	.92
	4	.84	.85	.85	.86	.87	.87	.88	.89	.89	.90	.90	.91
	8	.83	.84	.85	.85	.86	.87	.87	.88	.89	.89	.90	.90
	15–0	.83	.84	.84	.85	.86	.86	.87	.87	.88	.89	.89	.90
	4	.82	.83	.84	.84	.85	.86	.86	.87	.88	.88	.89	.89
	8	.82	.83	.83	.84	.85	.85	.86	.86	.87	.88	.88	.89
	16–0	.81	.82	.83	.84	.84	.85	.85	.86	.87	.87	.88	.88

MENTAL AGE

Chronological Age: Years	Months	11						12					
		0	2	4	6	8	10	0	2	4	6	8	10
	6–0	1.36	1.37	1.38	1.38	1.39	1.40	1.41	1.41	1.42	1.43	1.43	1.44
	4	1.32	1.32	1.33	1.34	1.34	1.35	1.36	1.37	1.37	1.38	1.39	1.39
	8	1.28	1.28	1.29	1.30	1.30	1.31	1.32	1.33	1.33	1.34	1.34	1.35
	7–0	1.24	1.25	1.25	1.26	1.27	1.28	1.28	1.29	1.29	1.30	1.31	1.31
	4	1.21	1.22	1.22	1.23	1.24	1.24	1.25	1.25	1.26	1.27	1.27	1.28
	8	1.18	1.19	1.19	1.20	1.21	1.21	1.22	1.22	1.23	1.24	1.24	1.25
	8–0	1.15	1.16	1.17	1.17	1.18	1.19	1.19	1.20	1.20	1.21	1.22	1.22
	4	1.13	1.14	1.14	1.15	1.16	1.16	1.17	1.17	1.18	1.19	1.19	1.20
	8	1.11	1.12	1.12	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17
	9–0	1.09	1.10	1.10	1.11	1.12	1.12	1.13	1.13	1.14	1.14	1.15	1.15
	4	1.07	1.08	1.08	1.09	1.10	1.10	1.11	1.11	1.12	1.12	1.13	1.13
	8	1.06	1.06	1.07	1.07	1.08	1.09	1.09	1.10	1.10	1.11	1.11	1.12
	10–0	1.04	1.05	1.05	1.06	1.06	1.07	1.07	1.08	1.09	1.09	1.10	1.10
	4	1.03	1.03	1.04	1.04	1.05	1.06	1.06	1.07	1.07	1.08	1.08	1.09
	8	1.01	1.02	1.03	1.03	1.04	1.04	1.05	0.05	1.06	1.06	1.07	1.07
	11–0	1	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.06
	4	.99	.99	1	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05
	8	.98	.98	.99	.99	1	1.01	1.01	1.02	1.02	1.03	1.03	1.04
	12–0	.97	.97	.98	.98	.99	.99	1	1.01	1.01	1.02	1.02	1.03
	4	.96	.97	.97	.98	.98	.98	.99	.99	1	1.01	1.01	1.01
	8	.95	.96	.96	.97	.97	.98	.98	.99	.99	.99	1	1.00
	13–0	.94	.95	.95	.96	.96	.97	.97	.98	.98	.99	.99	.99
	4	.93	.94	.94	.95	.95	.96	.96	.97	.97	.98	.98	.99
	8	.93	.93	.94	.94	.95	.95	.96	.96	.97	.97	.98	.98
	14–0	.92	.93	.93	.94	.94	.95	.95	.96	.96	.96	.97	.97
	4	.91	.92	.92	.93	.93	.94	.94	.95	.95	.96	.96	.97
	8	.91	.91	.92	.92	.93	.93	.94	.94	.94	.95	.95	.96
	15–0	.90	.91	.91	.92	.92	.93	.93	.94	.94	.95	.95	.96
	4	.90	.90	.91	.91	.92	.92	.93	.93	.94	.94	.95	.95
	8	.89	.90	.90	.91	.91	.92	.92	.93	.93	.94	.94	.94
	16–0	.89	.89	.90	.90	.91	.91	.92	.92	.93	.93	.93	.94

MENTAL AGE

Chronological Age: Years		13						14						15
	Months	0	2	4	6	8	10	0	2	4	6	8	10	0
	6–0	1.44	1.45	1.46	1.46	1.47	1.47	1.48	1.48	1.49	1.49	1.50	1.50	1.51
	4	1.40	1.41	1.41	1.42	1.42	1.43	1.43	1.44	1.44	1.45	1.45	1.45	1.46
	8	1.36	1.36	1.37	1.37	1.38	1.38	1.39	1.39	1.40	1.40	1.41	1.41	1.41
	7–0	1.32	1.32	1.33	1.33	1.34	1.34	1.35	1.35	1.36	1.36	1.37	1.37	1.38
	4	1.28	1.29	1.30	1.30	1.31	1.31	1.31	1.32	1.32	1.33	1.33	1.34	1.34
	8	1.25	1.26	1.26	1.27	1.27	1.28	1.28	1.29	1.29	1.29	1.30	1.30	1.31
	8–0	1.23	1.23	1.24	1.24	1.24	1.25	1.25	1.26	1.26	1.27	1.27	1.28	1.28
	4	1.20	1.21	1.21	1.22	1.22	1.22	1.23	1.23	1.24	1.24	2.24	1.25	1.25
	8	1.18	1.18	1.19	1.19	1.20	1.20	1.21	1.21	1.21	1.22	1.22	1.23	1.23
	9–0	1.16	1.16	1.17	1.17	1.18	1.18	1.19	1.19	1.19	1.20	1.20	1.21	1.21
	4	1.14	1.14	1.15	1.15	1.16	1.16	1.16	1.17	1.17	1.18	1.18	1.19	1.19
	8	1.12	1.13	1.13	1.14	1.14	1.14	1.15	1.15	1.16	1.16	1.16	1.17	1.17
	10–0	1.11	1.11	1.11	1.12	1.12	1.13	1.13	1.13	1.14	1.14	1.14	1.15	1.15
	4	1.09	1.10	1.10	1.10	1.11	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.14
	8	1.08	1.08	1.09	1.09	1.09	1.10	1.10	1.11	1.11	1.11	1.12	1.12	1.12
	11–0	1.06	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.10	1.11
	4	1.05	1.05	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.09	1.09	1.09	1.10
	8	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.08	1.08	1.08	1.09
	12–0	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07
	4	1.02	1.02	1.03	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06
	8	1.01	1.01	1.02	1.02	1.03	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.05
	13–0	1	1.00	1.01	1.01	1.02	1.02	1.02	1.03	1.03	1.03	1.04	1.04	1.04
	4	.99	1.00	1	1.00	1.01	1.01	1.01	1.02	1.02	1.03	1.03	1.03	1.03
	8	.98	.99	.99	1.00	1	1.00	1.01	1.01	1.01	1.02	1.02	1.02	1.03
	14–0	.98	.98	.99	.99	.99	1.00	1	1.00	1.01	1.01	1.01	1.02	1.02
	4	.97	.98	.98	.98	.99	.99	.99	1.00	1	1.00	1.01	1.01	1.01
	8	.97	.97	.97	.98	.98	.98	.99	.99	.99	1.00	1	1.00	1.01
	15–0	.96	.96	.97	.97	.97	.98	.98	.99	.99	.99	.90	1.00	1
	4	.95	.96	.96	.97	.97	.97	.98	.98	.98	.99	.99	.99	.99
	8	.95	.95	.96	.96	.96	.97	.97	.97	.98	.98	.98	.99	.99
	16–0	.94	.95	.95	.95	.96	.96	.96	.97	.97	.97	.98	.98	.98

A TEST FOR ART APPRECIATION

THEODORE F. KARWOSKI AND ERWIN O. CHRISTENSEN[1]

There is a need for objective tests in art. A test in art appreciation that functions should make it clear that art appreciation is a definite ability which may be measured, and that it is a phase of mental activity which cannot be identified with general intelligence.

The test, which is the subject of this article, was submitted to the art faculty of two universities and one art school. Although there was unanimous agreement on few questions only, nevertheless, a majority of these critics agreed, sufficient to warrant carrying the problem through.

The test combines three methods. The first and most satisfactory of these is the method of comparing two examples, a good one and a poor one. The answer is covered by five specific reasons. Of the five reasons only one is correct; four are wrong, do not apply, or fail to bring out the most important point. But all the reasons make sense and aim to be specific. It is of course, not always possible to do full justice to a work of art by summing up its qualities in a few words or in a brief sentence. For this reason only one of several possible right reasons was admitted as correct and care was taken that a competing reason would not be equally valid. The reasons were worded in simple and ordinary language so that the two could be used for elementary students.

The original purpose of introducing reasons in the test was methodological in character. It was desired that the subject taking the test would really prefer the examples chosen and not merely choose by chance. In comparisons of two pictures without reasons, the chance factor of 50 per cent would have made it necessary to include a very large number of comparisons in order to get a substantial score after deduction for chance had been made. By adding reason, chance was minimized with each reason.

In order to find out whether the inclusion of reasons affected the test in any other way, the same test was given to the same group, first without reasons and then with reasons. A comparison of the two methods indicated a high inter-correlation, but the method with reasons increased the difference between the highest and lowest scores.

[1] Reported at the annual meeting of the College Art Association in Chicago, Dec. 31, 1924.

To illustrate the method more clearly, the following example—Question No. 14 is quoted from the test.

A is better than B

because

B is better than A

1. The face is more sympathetic,
2. The features are clearer,
3. It is rich and subtle in effect,
4. It seems more life-like,
5. It is plainer and simpler.

This example compared a portrait by Chase with an ordinary photograph, and proved to be highly selective for different types of students. The phrasing of the reasons was adequate to bring out different responses. The four groups of subjects used, untrained in art, general art students, advanced art students, and art instructors chose the better picture regardless of reasons, in the following percentages: untrained 39, general art students 65, advanced art students 85, and art instructors 100 per cent. For the same example the percentage of correct choice with correct reason is as follows: 25 per cent for untrained; 50 for the general art students; 75 for advanced art students; and 88 per cent for art instructors.

Besides the method of comparison, above illustrated, the two other methods were: (1) judging a single picture in terms of good or not good, and checking one of the five reasons submitted with the question, and (2) selecting the best of four examples of similar subjects.

The test questions were 78 in number and involved a great variety of examples of art. Painting took up two-fifths of the test; architecture and sculpture together one-fifth; and the industrial arts, and abstract designs and color each one-fifth of the total.

This test was given to 505 persons of various levels of art training. Of these, 15 were instructors in art in high schools and 12 were college art instructors. The others were students from the universities of North Dakota, Minnesota, Chicago, and the Minneapolis School of Art. These people were divided into groups as follows: untrained students, general art students, advanced art students, and art instructors.

The advanced art group is perhaps the best from the point of view of definiteness in classification. It is composed of third year art students, and juniors and seniors in universities who majored in art.

The general art group is less definite in classification. It comprises students who had some art work in art school or college but not enough to be clsased as advanced art students. Thus some of these students approached those in the advanced art group in amount of art training, while others were nearer to the untrained group.

The untrained group is made up of students from the university of North Dakota who never had any art training, either in college or high school. North Dakota yields an ideal untrained in art student. Museums and centers of art are absent. The people of the state are homogeneous in the sense that there are no great differences between classes in respect to wealth and culture.

The data in Table I form a summary of the students tested, arranged in the groups described. The range of scores for each group is given, as well as the median, average scores and quartile.

TABLE I

	Number	Range	Median	Average	Quartile
Students with No Art Training:					
University of North Dakota.....	222	14–90	44	46	12
General Art Group:					
1. University of Chicago (summer session).....	22	38–104	61	64	
2. University of Minnesota.....	31	43–100	68	69	
3. University of North Dakota.....	67	34–100	63	63	
	120	34–105	64	64.6	13.5
Advanced Art Group:					
1. University of Chicago (summer session).....	39	40–120	72	74	
2. University of Minnesota.....	36	57–104	81	82	
3. University of North Dakota.....	24	48–118	95	94	
4. Minneapolis School of Art.....	37	41–116	81	83	
	136	40–120	81	82	13
Professional Art Group:					
1. Teachers of Art Schools of Minneapolis.....	15	65–128	97	95	
2. University Instructors.....	12	83–130	115	110	
	27	65–130	105	104	10
Total.....	505				

These data seem to indicate distinctly that the test selects students who have art appreciation and that art appreciation may be measured. The assumption on which the test is based is that those who have most art training are likely to have most art appreciation.

The above results may be plotted in curves. By plotting scores along the horizontal axis and the frequencies vertically the following curves of distribution of scores were derived:

In curves No. I, II, III, and IV, which represent the distribution of scores of the untrained, general, advanced, and professional art groups, the most impressive thing is the widespread overlapping between

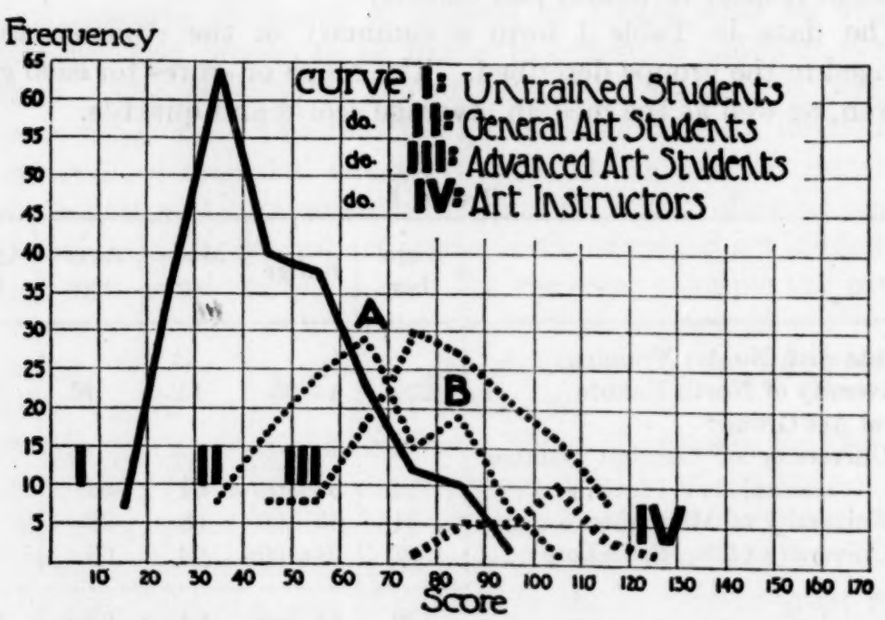


FIG. 1.—Curves showing distribution and frequency of scores for various groups.

groups. Of all the curves, No. I is the most skewed, indicating that the test is very difficult for the untrained students. On the other hand, curve No. IV is skewed toward the higher scores, indicating that the test is relatively easy for the professional group. Curves No. II and III overlap greatest and approach a normal curve of distribution. This type of distribution shows that the test is best adapted for these groups. In curve No. II there seem to be two modal points—one at A and the other at B—indicating the likelihood that this is a composite group, formed by two distinct types of students. As previously stated, this group is the least satisfactory in regard to classification. This is because it contains mostly students who had very little art, but includes some who had a course in the history of art, or one in art

appreciation. It may be that courses of this nature may be particularly helpful for the test. Still others (not many), had had considerable training but not quite enough to be classified as advanced art students.

In the main the groups are fairly distinct. The modal points of the groups are not separated by equal distances but that is probably due to lack of sufficient measures, because the average scores of the groups—44, 61, 81, and 100 are rather evenly distributed. With more cases the averaged and modal points would tend to coincide. More careful definition of the general art group would also tend to bring about such coincidence.

The overlapping of curves is quite typical but it is particularly interesting with reference to the untrained group, because this group is well defined in regard to lack of art training, not only as far as school work is concerned, but lack of artistic environment such as is presented in North Dakota. That students of this group should overlap into the highest group suggests that art taste is native or early acquired.

It was possible to compare the test scores of the individuals of the untrained group with the scores on the Thurstone Psychological Test No. 4 which most of the untrained students took on entrance at the university. The correlation between the two tests was worked out separately for 80 girls and 51 boys by Pearsons Product Moment formula. A slight positive correlation was found—.299 for girls and .264 for boys. The results are particularly interesting because it was felt that possibly the use of reasons would involve intelligence to such an extent as to show a high correlation with intelligence test marks. In that case there would hardly be a use for a test in art appreciation; an intelligence test would serve as well.

In constructing this test for art appreciation, it was intended to include as great a variety of art examples as would be practical for test purposes. Accordingly the test was divided into groups of questions. Part No. 1 consists of comparisons of paintings; part No. 2 is also a group of paintings, but in this case single pictures were judged rather than compared to another. Part No. 3 consists of examples from architecture and sculpture; part No. 4 was made up of examples from the industrial arts, and part No. 5 contains specially prepared designs intending to illustrate art in the abstract.

Due to this division of the test into five parts it is very easy to evaluate the response of the various groups of students to the five parts of the test. The following curves show the nature of the response of

the four classes of subjects tested. The vertical lines indicate the five parts of the test and on these are plotted the per cent of times the correct example, was chosen by the four groups of students. The curves are labelled I, II, III, and IV, and correspond to untrained, general, advanced, and professorial art groups.

The curves are interesting in that they tend to parallel each other. It would seem that the response to the test, by the four groups of subjects, was essentially similar in all respects except in quantity.

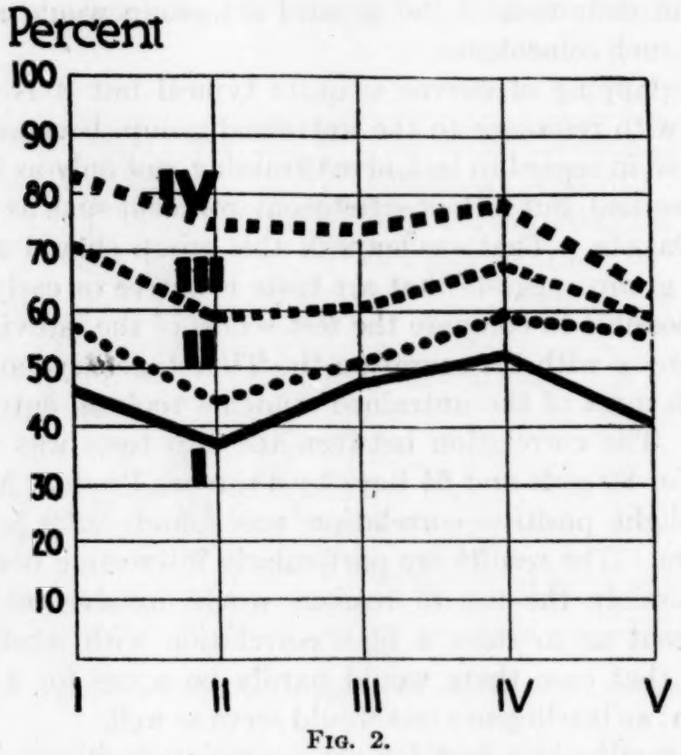


FIG. 2.

Hence the test is quantitative in its implications and the levels of art appreciation which the curves establish are levels of quantity rather than quality.

This question may be investigated more closely by means of correlations. For example, suppose we arrange the test questions in the order of difficulty for the untrained students, and correlate this order with the difficulty of the same pictures for the general art group, for the advanced art group, and the professional art group. This procedure may be taken for two sets of data—for correct choice of art example only and for correct choice plus correct reason.

Such correlations would reveal the extent to which the groups considered correspond in the order of difficulty of their responses to the questions of the test. The facts as indicated by Pearson's Product Moment formula are as follows:

Correlation between orders of difficulty between	Correct choice	Correct choice plus correct reason
Groups I and II	.874 ± .02 PE	.888 ± .02 PE
Groups I and III	.661 ± .05 PE	.793 ± .03 PE
Groups I and IV	.362 ± .08 PE	.491 ± .07 PE

These results indicate that in general the material that is usually preferred by the untrained group is also preferred by the other groups. There is a positive correlation in all cases, but the degree with which order of difficulty tends to correspond decreases progressively as we advance in the direction of the most advanced groups.

The method of correlation may be used to study the amount of correspondence between the order of difficulty for 'choice of correct example' and order of difficulty for 'choice of correct reason' for the same questions. In developing this test we had to assume, if we were to use reasons, that those who could appreciate art objects could also recognize why they appreciated these objects.

The following figures represent the correlation or correspondence between the 'choice of correct example' and 'choice of correct example plus correct reason' for the four groups of subjects.

Group I	.656 ± .05 PE
Group II	.656 ± .05 PE
Group III	.61 ± .06 PE
Group IV	.676 ± .05 PE

These correlations indicate that the tendency to score the correct reason in proportion to ease of appreciation of particular example of art, is almost identically the same for all the groups of subjects. Hence it is evident that if the reasons are carefully worked out choice of the better example of art will determine the choice of the correct reason, providing the subject is not guessing.

Conclusion

This investigation on art appreciation has proven that the method of testing appreciation by forcing the subject to give an opinion of

why one art example is preferred rather than another is suitable for test purposes. The method is an improvement on the ordinary impressionistic methods which merely call for expression of preference, in that it controls guessing. The reasons do not introduce factors not ordinarily employed in appreciation of art. The facts—that there is a high correlation between choice of correct picture and choice of correct reason, that the order of scores when test is given without reasons is similar to that when test is given with reasons, and that high intelligence does not necessarily give one an advantage when the reasons are used—indicates that the affective response is not changed by the introduction of reasons. On the other hand the use of reasons minimizes guessing, and therefore makes measurement of an individual's appreciation more valid.

AGE, SEX AND THE INTERRELATIONS OF MENTAL ATTAINMENTS OF COLLEGE STUDENTS[1]

JOHN W. GOWEN AND MARJORIE GOOCH

University of Maine

Categorically speaking, the history of centuries ago showed clearly that chronological age and mental age are by no means necessarily the same. Mental age as measured by mental ability is dependent on things of much greater fineness than is crude chronological age. Evident though this fact appears to us today, it was only after the work of Binet and others with one or another of the forms of mental tests that the true force of this conclusion appeared. These tests are frequently complicated. They are admittedly in the experimental stage, for the science of psychology itself is a young subject. Such being the case, is it any wonder that those whose life work does not bring them in contact with these tests are frequently heard to remark; "He is too young to have any ideas or to handle that work." Age in years is a fetter frequently binding the mentally able to mediocre accomplishment. Under these circumstances, it is of some importance to examine the relation that chronological age bears to mental ability in as many lines as possible. The mental ability as shown by the ranks of college students furnishes a good field for investigation of this subject. While the obvious applicability of the data to this and other problems would seem to make it likely that they would have been extensively explored, the facts of the case are that the field has been quite largely neglected.

Like chronological age, sex furnishes another variable which has been supposed to play a large part in the psychic expression of the individual. A not inconsiderable mass of evidence has been collected on this polyphased problem. The method of approach has again been largely through the agency of the mental test and has not dwelt on the more complicated but perhaps more practical problem of the effect of sex on the whole education of the individual. The analysis of the influence of these two variables, age and sex, on the mental ability of the individual as displayed by their college work, form the theme of this paper. Following this preliminary analysis the correlations exist-

[1] This is the second article of a series by these authors. The first appeared in the *Journal of Educational Psychology* for November, 1925, entitled "The Mental Attainments of College Students in Relation to Previous Training."

ing between mental attainments in the first year of college will be studied.

Methods

Our results are based on the college and high school records of 927 graduates of the University of Maine. The method of gathering the data and the statistical treatment are more fully explained in the preceding paper (*loc. cit.*). Suffice it to say, that the figures are founded on those college graduates of the University who entered between 1909 and 1917 inclusive.

Influence of Age on Mental Attainments

Our data show that the entrance ages range from 16 to 28 years, with the average at 19.5. Table I gives the correlation of age with the different college students.

Table I.—The Relation of Age to College Rank

	Average age	SD	Correlation coefficient
English	19.5 ± .033	1.47 ± .024	−.048 ± .023
Chemistry	19.5 ± .038	1.50 ± .027	+.022 ± .025
Algebra	19.4 ± .038	1.38 ± .027	+.065 ± .027
Analytical geometry	19.4 ± .040	1.36 ± .028	+.089 ± .029
Elementary French	19.4 ± .119	1.42 ± .084	+.042 ± .083
Advanced French	19.5 ± .077	1.60 ± .054	+.013 ± .048
German	19.5 ± .052	1.46 ± .037	−.016 ± .036

The correlations for all seven of the subjects—English, chemistry, algebra, analytical geometry, elementary French, advanced French, and German—are small. Four out of the seven—chemistry, elementary French, advanced French, and German—are exceeded by their probable errors. Only one, that of analytical geometry is at all significant. This correlation, .089 ± .029, is three times the probable error. Two of these correlations, English and German, are negative, showing the younger student to be the better.

The insignificant correlations obtained from our data would certainly suggest that grades and actual chronological age are not very closely related. Innate mental ability as an attribute of the individual appears to be the greater force in determining mental attainments.

Unfortunately there is a loophole even to this conclusion for the lack of correlation of mentality with age can be explained by the assumption that the younger college student has a greater mentality than the older student, and that the older student has equalled the attainments of the younger because of his experience. The reasons for the older students being in the same class with those right out of high school are, of course, many and diverse. It is undoubtedly true in some cases that lack of finances is the reason for the delay in college education. For such cases the above hypothesis is borne out in part by the work of Terman. Terman[1] divided 492 of the school children to whom he gave mental tests into groups according to the social class of their families. He found (p. 72) that the median intelligence quotient for children of the superior social class is seven points above the median for those of the average social position. By the age of 14, this means that children of superior classes are one year above the average.

The problem is undoubtedly complex. However, insofar as it concerns education or similar tests of ability it is evident that chronological age is an unreliable criterion of what the mental ability of the individual may be. In our data it is evidently a factor of no importance to the results deduced. No further consideration will be given to chronological age.

Influence of Sex on Mental Attainments of Students

The coeducational college offers an excellent testing ground for the influence of sex on the mental attainments of the college students. Here all students are on equal footing. They are in the same classes, they are tested by the same standards and under the same conditions. Such an environment presents a real opportunity for analyzing any difference in ability which may exist in the two sexes.

While the psychological experiments on the difference of the sexes have generally arrived at the conclusion that women are the equals of men in mental ability, they do not necessarily mean that the mental attainments of the sexes are the same. Gates[2] finds women better in memory than men, while the latter are better in reasoning but by a smaller difference. Probably one of the most important experiments in helping women substantiate their claim of equality in potential

[1] Terman, Lewis M.: "The Measurement of Intelligence." 1916.

[2] Gates, A. J.: Experiments on the Relative Efficiency of Men and Women. *Psychological Review*, Vol. XXIV, p. 139.

mental ability[1] is that of Mall on the character of the brain of the two sexes. Mall found no differences sufficient to enable an unbiased observer to distinguish the sexes by the appearance of the brain. Thorndike's work[2] leads him to the conclusion that the sexes are essentially equal in mental capacity, each sex offering almost as great differences *inter se*.

If women, as students, were the equals of men in every respect we should expect to find the average ability of each to be the same. Table II shows the women to have significantly higher averages than do the men.

TABLE II.—THE AVERAGE COLLEGE RANKS OF THE SEXES

	English	Chemistry	Algebra	Analytical geometry	Elementary French	Advanced French	German
Men.....	79.9±.3	80.3±.3	81.6± .3	81.1± .3	78.7±1.1	81.4±.6	81.4±.4
Women...	85.0±.4	79.8±.7	85.7±1.2	83.4±1.2	90.1± .8	91.1±.4	89.9±.7

In every case except chemistry, women are markedly better students. In elementary French, women are 11.4 ahead: in advanced French this difference is 9.7 in the women's favor, in German it is 8.5, and in English 5.1. In the mathematics courses the difference is not so marked, being 4.1 and 2.3 in algebra and analytical geometry, respectively. If we can assume mathematics and chemistry to be better indicators of reasoning ability, which is the popular view, and languages the better indices of memory, then we have further evidence supporting Gates' work. In this work, to discount the factor of preparation in which women are supposed to exceed men, he read a newspaper clipping containing several details to a class of men and women, simply with the admonition to pay close attention to the reading. At the end of the class, and again several weeks later, without warning, questions were asked as to the details of the clipping. In both these tests, the women remembered more points correctly.

If we could discount the factor of selection (which our evidence does not seem to allow us to do but which Gates seems to do), and not have preparation as a factor, we would have pretty strong evidence that women were not only men's equal but their superiors in mental

[1] Wooley, H. T.: A Review of the Recent Literature on the Psychology of Sex. *Psychological Bulletin*, Vol. VII, p. 335 also article on same subject in *Psychological Bulletin* Vol. XI, p. 353.

[2] Thorndike, E. L.: *Educational Psychology*. (1910) Revised edition, pp. 248.

ability. The college grades from which our data are compiled do not deduct the factor of preparation. So that the differences in college grades are not conclusive evidence of the position of women in actual ability. Furthermore they do not take into account the possible factor of the selection of only the best women as college students. If this selection of students is more stringent for women than for men we would expect the same population to differ by larger amounts in their previous high school work. Table III shows the averages in high school.

TABLE III.—AVERAGE HIGH SCHOOL RANKS OF THE SEXES

	English	Chemistry	Algebra	Geometry	Latin	Elementary French	Advanced French
Men.....	83.2±.2	84.5±.3	84.0±.3	84.5±.3	83.2±.3	81.8±.3	82.7±.5
Women..	89.1±.4	86.9±.7	88.0±.5	85.7±.5	87.5±.5	87.7±.4	87.1±.5

Table III shows a different range of differences. In high school these same women were higher than the men in every case: 5.9 in English and elementary French, 4.3 in Latin, 4.4 in advanced French, 2.4 in chemistry, 4.0 in algebra, 1.2 in plane geometry.

The conclusion that those women who go to college are slightly better than the men in their mental attainments in high school is obvious. This problem is pushed back to the more embryonic stage for its final solution. However important it might be, we are not at present in a position to further analyze this problem through the grammar school and into the home back to the cradle. We cannot differentiate between two possible conclusions; (1) that women are slightly better equipped to attain success as measured by their mental attainments in high school and college or that (2) the women who go to college as a group are more heavily selected on a basis of their ability to achieve success in high school and college than are the men who attend college. Without tracing the circuitous route followed through the literature necessary to arrive at a working theory it appears probable that both factors play a part but that the major factor is that of a more stringent selection among women entering college.

Interesting as this problem is, it is in a sense a side line to that now before us. No matter what mechanism brought about these differences, if they are compared with the differences which exist in the college ranks between men and women, it becomes clear that this selection is not the only factor, for the differences for college women are consistent

in showing the women slightly superior to the men even though the more stringent selection of high school women for college be discounted. In view of these differences between the two sexes, the question naturally arises as to whether the differences also extend to the relationships between high school and college subjects.

Table IV shows the correlation coefficients of the women between such college and high school subjects as have large enough numbers to make them of statistical value. Table V shows similar relations for the men students.

TABLE IV.—CORRELATION COEFFICIENTS FOR HIGH SCHOOL AND COLLEGE SUBJECTS AS OBTAINED FOR WOMEN STUDENTS

College	High school					
	English	Chemistry	Algebra	Geometry	Latin	Elementary French
English	.309 ± .055	.250 ± .088	.166 ± .061	.267 ± .058	.256 ± .059	.120 ± .063
Chemistry	.395 ± .072	.547 ± .082	.217 ± .084	.316 ± .079	.411 ± .075	.283 ± .083
Algebra	.161 ± .143					
Analytical geometry	.361 ± .128					
Advanced French	.172 ± .078	.309 ± .106	.066 ± .082	.069 ± .081	.184 ± .081	.311 ± .073
Elementary German	.130 ± .133					

TABLE V.—CORRELATION COEFFICIENTS FOR HIGH SCHOOL AND COLLEGE SUBJECTS AS OBTAINED FOR THE MEN

College	High school					
	English	Chemistry	Algebra	Geometry	Latin	Elementary French
English	.233 ± .030	.180 ± .035	.152 ± .030	.167 ± .032	.176 ± .037	.245 ± .032
Chemistry	.202 ± .032	.168 ± .037	.223 ± .032	.221 ± .032	.222 ± .039	.310 ± .033
Algebra	.212 ± .035	.333 ± .037	.416 ± .031	.421 ± .031	.254 ± .041	.409 ± .033
Analytical geometry	.268 ± .035	.300 ± .039	.378 ± .033	.340 ± .034	.281 ± .042	.375 ± .035
Advanced French	.174 ± .086	.042 ± .115	.148 ± .089	−.031 ± .088	.121 ± .106	.035 ± .095
Elementary German	.173 ± .045	.249 ± .047	.294 ± .043	.282 ± .044	.383 ± .046	.349 ± .042

Examination of the correlation coefficients in Tables IV and V shows that while small in actual amount, a significant relation exists between the high school ability of either men or women and their ability in college. The average correlation coefficient for the women is .252 and that for the men .173. While such a difference is probably not significant in a statistical sense, it may indicate a biological or

sociological difference between the sexes. It is conceivable that the high school work of women should be a better measure of their future accomplishments in college because of the admitted fact that women spend more time in study than men, thus tending to differentiate their mental capacities more accurately than would be the case for men. Such a possibility is further strengthened by the size of the differences of the homologous correlation coefficients. Of the 21 combinations of subjects compared, 14 show a larger relationship between the high school and college ranks in subjects taken by the women. Of the large differences five favor the women and one the men. On the other hand the probable errors of the differences are large. Only one difference is more than three times the probable error of this difference. In view of these facts it seems probable that the men and women are not differentiated to any marked degree in their ability or in the accuracy with which their college attainments may be predicted from their high school accomplishments. The conclusion is also borne out by what is to follow.

No evidence for specialized ability, *e.g.*, mathematical genius *vs.* relatively mediocre attainments in other fields, appears in this study. The system of high school and college teaching seems rather to encourage generalized ability in all lines and to stifle brilliance in a single field. With these facts in mind, the problem of the interrelation of the different subjects in college may be considered. Here the environment is more homogenous than that of the 300 different high schools, and presumably the grade of teaching is different.

College Attainments in Their Relation to Each Other

The intercorrelations between the subjects which are taken during the same year in college are shown in Table VI.

Table VI.—Relation between College Grades

Subject	Mean	SD	r	Subject	Mean	SD	n
English....	80.4 ± .3	9.88 ± .18	.386 ± .021	Chemistry.....	80.3 ± .2	9.68 ± .17	708
English....	80.3 ± .3	9.79 ± .19	.301 ± .026	Algebra.....	82.1 ± .3	11.80 ± .23	588
English....	80.5 ± .3	9.87 ± .20	.328 ± .026	Analytical geometry....	81.6 ± .3	10.86 ± .22	544
English....	80.6 ± 1.0	12.20 ± .73	.440 ± .066	Elementary French.....	82.9 ± .9	10.97 ± .65	64
English....	82.6 ± .5	9.48 ± .32	.336 ± .042	Advanced French.....	85.9 ± .4	9.26 ± .31	199
English....	80.5 ± .3	9.78 ± .25	.338 ± .032	German.....	82.6 ± .3	9.59 ± .24	358

The difference between these correlations is even less significant than those with high school ranks. This is a reasonable consequence of a more uniform marking system in the University as compared with

that of the 300 high schools. Here, too, we find the correlation of English with chemistry slightly higher than with advanced French or German. These differences are not significant, but they are interesting as indicating that ability in languages need not be distinct from ability in sciences.

TABLE VII.—RELATION BETWEEN COLLEGE GRADES

Subject	Mean	SD	r	Subject	Mean	SD	n
Chemistry...	80.8 ± .3	9.91 ± .20	.294 ± .028	Algebra.....	82.3 ± .3	11.51 ± .24	514
Chemistry...	80.8 ± .3	9.96 ± .22	.374 ± .026	Analytical geometry..	81.4 ± .3	10.97 ± .24	479
Chemistry...	78.6 ± 1.2	11.11 ± .84	.479 ± .080	Elementary French...	78.4 ± 1.1	10.73 ± .81	40
Chemistry...	79.2 ± .7	10.68 ± .47	.343 ± .041	Advanced French....	83.9 ± .6	9.79 ± .43	118
Chemistry...	81.0 ± .4	9.46 ± .25	.334 ± .033	German.....	82.0 ± .4	9.64 ± .26	319

The relation between chemistry and the other subjects, as given in Table VII, shows the same general characteristics. The higher correlation between chemistry and elementary French is really not significant because of its higher probable error, due to small numbers of individuals taking these courses.

The correlations between the mathematics courses would probably most nearly approximate that of Rietz[1] on the relations between mathe-

TABLE VIII.—RELATION BETWEEN COLLEGE GRADES

	Mean	SD	r		Mean	SD	n
Algebra....	82.8 ± .3	11.27 ± .24	.583 ± .020	Analytical geometry..	81.4 ± .3	11.33 ± .24	509
Algebra....	81.2 ± 1.3	12.62 ± .93	.384 ± .087	Elementary French...	81.4 ± 1.7	10.28 ± .76	42
Algebra....	83.1 ± .8	10.93 ± .58	.267 ± .068	Advanced French....	83.5 ± .7	10.06 ± .53	82
Algebra....	82.1 ± .5	11.92 ± .35	.278 ± .038	German.....	82.0 ± .4	9.82 ± .29	267
Analytical geometry	80.1 ± 1.4	13.95 ± 1.0	.319 ± .091	Elementary French...	80.0 ± 1.2	11.54 ± .84	43
	79.7 ± .9	12.70 ± .65	.411 ± .060	Advanced French....	83.4 ± .7	9.70 ± .49	88
	82.3 ± .4	10.07 ± .29	.368 ± .035	German.....	82.0 ± .4	9.89 ± .29	274

matics and law in Dartmouth students. The combined data of five years totalling 161 students gave a correlation coefficient of .528 ± .040. From this the natural conclusion is drawn that students high in mathematics tend to be higher in law than the average student. Mathematics and law are assumed to be subjects largely dependent on different types of reasoning. The nearest approach in our data to differing types of reasoning in the freshman year, is that displayed in

[1] Rietz, H. L.: On the Correlations in the Marks in Mathematics and Law. *Journal of Educational Psychology*, Vol. VII, p. 87.

algebra and analytical geometry. If this is at all a legitimate comparison and we should find a correlation as large as Rietz did, we should be justified in believing that reasoning ability was transferable. Our correlation between these subjects, algebra and analytical geometry, based on the 509 students available, is .583 ± .020.

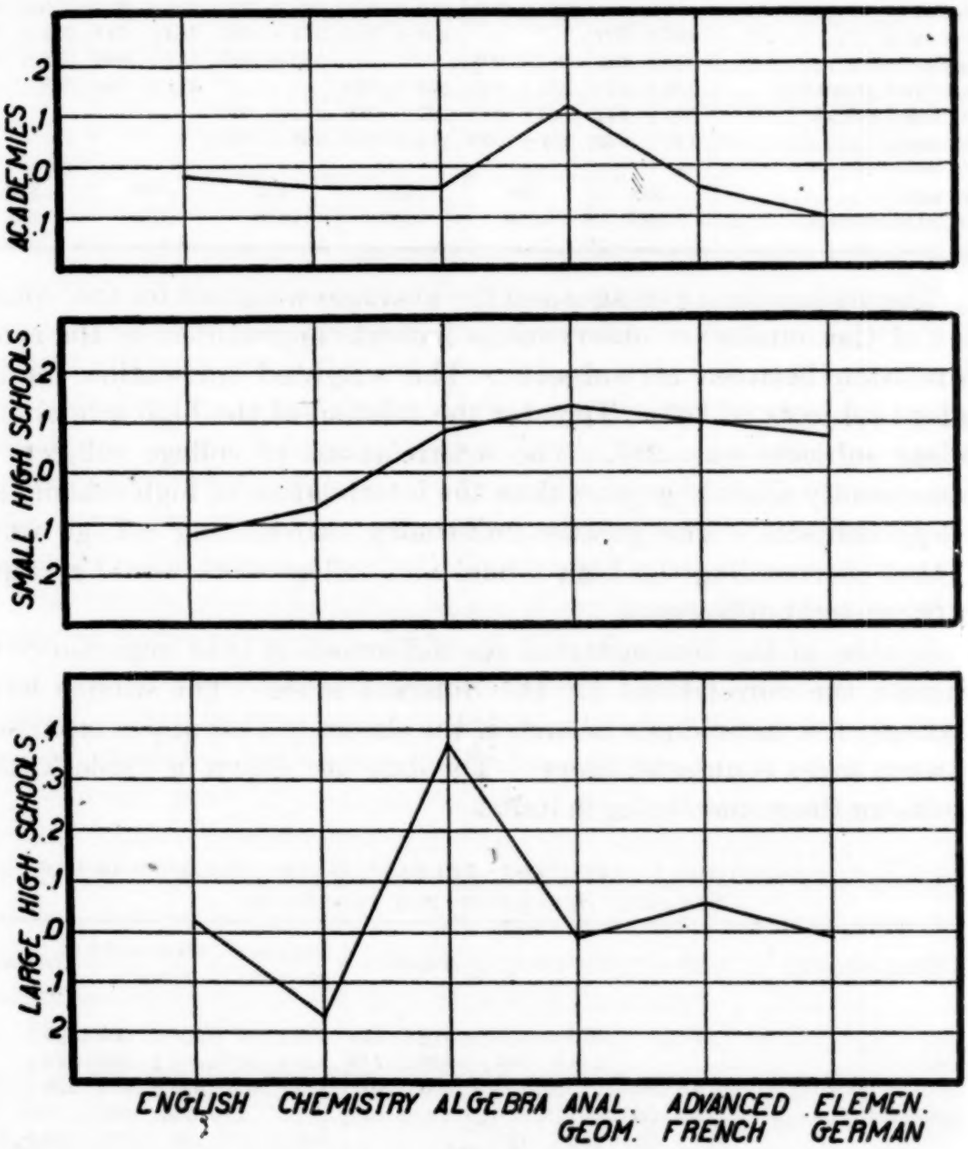


FIG. 1.

The results for this study of the interrelation of attainments in college work are shown in Fig. 1. The complete data showing the correlation coefficients between all the subjects for the combined sexes are shown in Table IX.

TABLE IX.—CORRELATION COEFFICIENTS SHOWING THE RELATION BETWEEN THE MENTAL ATTAINMENTS OF COLLEGE STUDENTS. COLLEGE SUBJECTS COMPARED

College subject	English	Chemistry	Algebra	Analytical geometry	Advanced French	German
English	.....	.386 ± .021	.301 ± .026	.328 ± .026	.336 ± .042	.338 ± .032
Chemistry	.386 ± .021	.....	.294 ± .028	.374 ± .026	.343 ± .041	.334 ± .033
Algebra	.301 ± .026	.294 ± .028	.....	.583 ± .020	.267 ± .068	.278 ± .038
Analytical geometry	.328 ± .026	.374 ± .026	.583 ± .020	.....	.411 ± .060	.368 ± .035
Advanced French	.336 ± .042	.343 ± .041	.267 ± .068	.411 ± .060		
German	.338 ± .032	.334 ± .033	.278 ± .038	.368 ± .035		
Average	.338	.346	.345	.413	.339	.330
Weighted average	.346	.356	.362	.410	.339	.331

The arithmetical averages and the averages weighted for the square root of the number of observations indicate approximately the same correlation between all subjects. The weighted correlation for all college subjects is .363. That for the relation of the high school and college subjects was .283. The interrelations of college subjects is consequently slightly greater than the interrelation of high school and college subjects. The greater uniformity surrounding college work *vs.* that surrounding the high school and college work would account for these slight differences.

In view of the demonstrated sex differences it is of importance to examine the correlations for the different sexes. The women have relatively few individuals in each of the classes so that any comparison between sexes is unsatisfactory. The data are shown in Table X, the results for the women being in italics.

TABLE X.—CORRELATION COEFFICIENTS BETWEEN RANKS OBTAINED IN COLLEGE SUBJECTS SEPARATED FOR THE SEXES

College subject	English	Chemistry	Algebra	Analytical geometry	Advanced French	German
		.379 ± .023	.319 ± .026	.320 ± .027	.198 ± .063	.295 ± .034
English	.....	*.415 ± .065*	*−.040 ± .119*	*.121 ± .117*	*.294 ± .063*	*.575 ± .076*
	.379 ± .023	.....	.296 ± .027	.373 ± .027	.373 ± .064	.328 ± .035
Chemistry	*.415 ± .065*	.....	.....	.....	*.113 ± .109*	
	.319 ± .026	.296 ± .027	.....	.568 ± .020	.245 ± .073	.316 ± .036
Algebra	*−.040 ± .119*	.....	.....	*.487 ± .091*		
	.320 ± .027	.373 ± .027	.568 ± .020	.....	.437 ± .066	.378 ± .035
Analytical geometry	*.121 ± .117*	.....	*.487 ± .091*			
	.198 ± .063	.373 ± .064	.245 ± .073	.437 ± .066	.....	too few
Advanced French	*.294 ± .063*	*.113 ± .109*				
	.295 ± .034	.328 ± .035	.316 ± .036	.378 ± .035	too few	
German	*.575 ± .076*					

Table X shows that the results for the men are essentially the same as the results for the whole group. The results for the women are scattering due to lack of numbers. The average weighted correlation coefficient for the men was .352; that for the women was .295. It will be noted that this comparison favors the men while the correlation for high school and college attainments favors the women. The lack of numbers makes the comparison of no significance, however.

Considering the results as a whole, the analysis of the data for the college work supports the same views as those expressed for the relation between the high school and college work.

1. Mental ability appears to be an attribute of the individual.

2. The expression of this attribute in high school and college work is approximately uniform for all subjects studied. As pointed out, it seems altogether probable that in these college and high school tests of mental ability a condition exists which is quite different from that of the tests for the mental ability of the genius as used in inheritance and sociology. Our educational system tends toward a test for general mental ability rather than a test for ability in any one line of endeavor. The test is then one for basal intelligence and possible training in all subjects rather than a test of mental ability in a specific subject. If this be the case the explanation of the general uniformity of the relations between mental ability of a student in all subjects appears more likely to be some such general factor as pride in the work accomplished. This interpretation is not in conflict with the inheritance of specific mental abilities but rather in line with it.

Of even more importance to the individual and perhaps to the teacher is the establishment of the fact that mental success is an attribute of that individual and, whether it be measured by high school or college ranks, the indications are that in either case the individuals high in one subject are more likely to be high in the others and the students low in one are likely to be low in the others.

It is difficult to interpret these results in terms of previous training, either by the transfer of identical elements or by the transfer of reasoning power. In the first place, if the identity were of subject-matter and were strongly operative, a decided correlation between similar subjects and a less correlation between unlike subjects would be expected. No such relationships between the different subjects exists. On the other hand, if the transfer were one of reasoning power or ability to learn, the correlation between all subjects would be about the same—a result which is actually obtained; but at the same time we

should expect that the quality of the training received in high school would be markedly shown. As will be shown in the third paper this result is not obtained. The quality or quantity of the high school training has not differentiated the students to any appreciable extent. The answer appears to be, rather, that mental ability is an attribute of the individual, probably governed by inheritance and possibly influenced in its expression by the home in which the student was brought up.

Summary

There is no direct correlation between mental ability and chronological age. In fact, the correlations of age of entrance with college English and German are both negative (−.048 and −.016 respectively) giving what slight advantage there is to the younger students. The correlations with age and the other subjects range from .013 with advanced French, .022 with chemistry, .042 with elementary French, to 0.65 with algebra, and .089 with analytical geometry. As compared with the probable errors, only the correlation of age with analytical geometry is significant. The results of this practical test in education check with those of other investigators using a form of the mental test in showing that mental age and chronological age are by no means synonymous. Furthermore, it is clear that the variable age plays little or no part in the effect of previous training on accomplishment in college work.

In conformity with the work of other investigators using a somewhat different method, women were shown to be equal if not slightly superior to men in their mental attainments in college. Women's superiority is slightly greater in the languages than in the sciences. For example in English, women lead by 5.1, in elementary French by 11.4, in advanced French by 9.7, in German by 8.5, while in chemistry men lead by .5. In algebra women are again ahead by only 4.1, and in analytical geometry by only 2.3.

The study of the high school ranks indicates that women who go to college are a selected group, much better in their high school attainment than the men who go to college. With this fact in mind it is not surprising to find the women in college higher than the men in their college work. The differences between sexes is not so great in college as in high school.

The analysis for the separate sexes of the records for the relation of high school and college work and for the attainments of the students

in college work showed that mental ability behaves as an attribute of the individual expressed by this individual uniformly in all of the subjects studied. Any tendency to genius in one line is conspicuous by its absence. Whether this absence is due to the dead level set by the requirement that any and all students must pass a certain grade of work in all subjects or to the lack of sufficient stimulus in college to prod the individual to the effort necessary to the accomplishments of the genius in any one line is not known. In view of the observed facts the stimulus for college attainment, if there is one, appears to be generalized to cover ability in all subjects.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY C. O. MATHEWS
Graduate Student, Teachers College
Columbia University

Intelligence Testing

Performance Tests for Children of Pre-school Age. Rachel Stutsman. Genetic Psychology Monographs, Vol. 1, No. 1, Jan., 1926. This is a preliminary report of the tests standardized by the Merrill Palmer School for which norms have been determined.

Intelligence among Normal School Students. Stephen G. Rich and Charles E. Skinner. Educational Administration and Supervision, Dec., 1925, 639–644. Data is given showing the standing of normal school students on several standardized tests.

An IQ Study. Frederick M. Foster. California Quarterly of Secondary Education, Oct., 1925, 73–74. This is a comparative study of 249 students of ten nationalities on the Terman group test.

An Analysis of Language Factors in Intelligence Tests. Dorothy W. Seago. Mental Measurement Monographs, No. 1. This study was made in order to ascertain "the nature of language factors and the extent to which they function in intelligence tests."

A School Intelligence Test. S. Bowie and A. R. Laws. The Forum of Education, June, 1925, 134–140. The results of two British tests, the Northumberland and Simplex, are used to point out some of the difficulties in comparing pupils by their intelligence scores.

The Intelligence Quotient as a Prognosis of Success in Physics. A. W. Hurd. The School Review, Feb., 1926, 123–128. The correlation between the average IQ obtained from two group tests and average scores on physics tests is $76 \pm .03$.

A Study of the Intelligence and Achievement of the June, 1925, Graduating Class of the Grover Cleveland High School, St. Louis. 1. W. D. Shewman. School Review, Feb., 1926, 137–146. Data is given relative to the relationship between intelligence and achievement in the various subjects for several hundred cases.

Paternal Occupations and Intelligence of Offspring. Peter Sandiford. School and Society, Jan. 23, 1926. The results are given of a survey of high school, normal school and university populations of British Columbia.

Predicting Success in College at the Time of Entrance. J. B. Johnston. School and Society, Jan. 16, 1926, 82–88. This is a report of the performance of groups selected at the University of Minnesota over a period of several years.

Achievement Testing

The Use of Standardized Educational Tests in School Surveys. Nell Bomar Johnston. Educational Administration and Supervision, Dec., 1925, 599–607.

All facts concerning the use of standard achievement tests are traced through the reports of 25 school surveys.

The Effect of Objective Standards upon Composition Teachers' Judgments. Earl Hudelson. Journal of Educational Research, Dec., 1925, 329–340. The author shows the speed of progress with which teachers' judgments become more reliable by the use of an objective scale.

A Study of Different Types of Teachers' Tests. Charles E. Ozanne. The School Review, Jan., 1926, 42–53. Pupils' preferences and opinions are given with regard to completion, essay and reasoning tests.

Sectioning Junior High School Pupils by Tests and School Marks. Fowler D. Brooks. Journal of Educational Research, Dec., 1925, 359–369. Numerous criteria for sectioning junior high school pupils are evaluated and compared.

Minimum Essentials in Reporting Data on Standard Tests. G. M. Ruch. Journal of Educational Research, Dec., 1925, 349–358. The author discusses criteria for evaluating tests and measurements with special emphasis upon methods of reporting reliabilities.

A Selected and Annotated Bibliography on Standardized Tests for Secondary Schools. Frank C. Touton. California Quarterly of Secondary Education, Jan., 1926, 253–271. The tests are classified according to subjects. The publisher, cost, grades where useful, time limits and purpose are given for each entry.

Building a Scale of Purely Composition Quality. S. A. Leonard. The English Journal, Dec., 1925, 760–775. This is a scale based on the grades of the Nassau County Supplement to the Hillegas Scale after errors in mechanics had been eliminated.

Suggestions for the Training of Teachers in the Use of Educational Measurements. A. R. Mead. Educational Administration and Supervision, Jan., 1926, 23–43. These suggestions relate to the objectives, content, methods of teaching and measuring the results of such a course.

Achievement of California High School Seniors in Reading Comprehension, Vocabulary and Spelling. Frank C. Touton and Karl K. Heilman. California Quarterly of Secondary Education, Jan., 1926, 175–194. This is a report of a survey of the reading, vocabulary and spelling abilities of California high school seniors.

Psychology of School Subjects

The Development of Number Concept in Children of Pre-school and Kindergarten Ages. Harl R. Douglass. Journal of Experimental Psychology, Dec., 1925, 443–470. This development is measured for concepts of numbers 1 to 10 in children of ages from 4 years 6 months to 6 years.

What Words Should Children Be Taught to Spell? III. Limitations of the Adult Standard of Selection. Frederick S. Breed. The Elementary School Journal, Dec., 1925, 292–306. Composite lists of words from childhood composition vocabularies and adult correspondence vocabularies are compared, and a detailed analysis made of words used by children only, by adults only and words common to both.

An Experimental Study of the Merits of Extensive and Intensive Reading in the Social Sciences. Carter V. Good. School and Society, Dec. 12, 1925, 758–760.

On the whole, extensive reading seems superior to intensive according to this study.

Number Needs on Children's Reading Activities. Clara Martin Partridge. The Elementary School Journal, Jan., 1926, 357–366. This is a presentation of the number situations occurring in the text books of California and in the periodicals read by public school children.

The Effect of Summer Vacation on Children's Mental Ability and on Their Retention of Arithmetic and Reading. Mildred V. W. Patterson. Education, Dec., 1925, 222–228. Results on standardized tests show that these children gained in intelligence, remained about constant in reading ability and lost in arithmetic ability during the summer vacation.

What Books Fit What Children. Carleton Washburne and Mable Vogel. School and Society, Jan. 2, 1926, 22–24. The building of a graded book list by experimental methods is reported.

Character and Personality

Bibliography on Trait Measurements and Their Practical Use. G. V. Whaley. California Quarterly of Secondary Education, Jan., 1926, 236. A short bibliography of eleven articles and monographs.

Why Intelligence Scores Are Not More Highly Predictive of School Success. W. Hardin Hughes. Educational Administration and Supervision, Jan., 1926, 44–48. Exceptional cases of the lack of correlation between personality traits and intelligence are set forth and discussed.

Personality Traits and the College Success of High School Graduates. W. Hardin Hughes. California Quarterly of Secondary Education, Jan., 1926, 225–236. Data is presented showing the relationship between personality traits and academic success.

Some Preliminary Results on the Nature of Developmental Age. Paul Hanly Furfey. School and Society, Feb. 6, 1926, 183–184. Developmental age is used to designate "the gradually increasing maturity of the child's personality reactions." This is a study of relationships.

Some Social, Age and Sex Differences Shown in Children's Choice of Ideals. Eva Macaulay. The Forum of Education, June, 1925, 105–114. This is an attempt to study ideals by an analysis of 1600 children's responses to "Whom would you most wish to be like?" and "Why?"

Miscellaneous

Some Experimental Studies of the Results of College Note-taking. C. C. Crawford. Journal of Educational Research, Dec., 1925, 379–386. The effect of note-taking upon immediate and delayed recall as tested by the essay and true-false types of tests is shown.

Pupil Reactions to School Reports II. W. A. Barton, Jr. The School Review, Jan., 1926, 42–53. Opinion of pupils are reported on the items, form and frequency of reports and sex differences in these opinions are shown.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY JOHN HOCKETT
The Lincoln School of Teachers College

Reading Comprehension in the High School

Comprehension Difficulties of Ninth Grade Students in the Study of Literature, by Theo. W. H. Irion. Teachers College, Columbia University, 1925. Pp. V + 116.

How well do ninth grade pupils, from their own reading solely, comprehend typical selections from the literature commonly assigned to them? The above dissertation offers some very interesting and significant information for the answer of this question. Dr. Irion gave four carefully devised literary comprehension tests to 202 ninth grade pupils in a New York City high school. In each test the pupils were required to read a selection and then, through multiple choice exercises, indicate their understanding of it with respect to conclusions, important points, unusual expressions, unusual pivotal words, and related general information. Three of these selections were typical of the literature found in ninth grade courses of study,—an extract from Cooper's *The Spy*, Byron's *The Destruction of Sennacherib*, and a part of Act I, Scene 3, of Shakespeare's *Julius Caesar*. The fourth selection was exposition, straightforward and unadorned—a paragraph from Darwin's *Origin of Species*, the author's name not mentioned. The median percentage scores on these tests were, respectively 56, 43, 43, 57, giving an average comprehension score of 50. The most important conclusions drawn by the author with respect to literary comprehension are (1) that it varies markedly with different types of literature, (2) that it correlates to a considerable degree with the Thorndike-McCall Reading scores, (3) that it is not sufficient in pupils with IQ's below 100 to make study of these selections profitable, and (4) that it cannot be assumed for most ninth grade pupils on the basis of their own reading.

The practical implications for the teaching of literature are readily apparent. Teachers should realize how much more they know about a classic than do their pupils; they should be much more concerned about

the mere reading ability of their pupils; they should make increasing provision for individual differences. The whole study is an unusually good one, both as courageous research in the muddled field of English and as suggestive technique for further investigation. Necessarily, there are a number of subjective elements affecting the procedure, such as the initial choice of the literary selections, the content of the test exercises, and the actual giving of the tests. These, however, are fully recognized by the author, who at the same time almost importunes us to believe that in these instances he has used every means in his power and experience to validate his judgment. Statistically the work is excellent. Distributions, variabilities, central tendencies, correlations, probable errors, and formulas are all presented opportunely and correctly. And, what is a good deal more, their meanings appear to be comprehnded. This, in the case of one whose main interest is English Literature, is worth mentioning.

M. H. Willing.

Additional Evidence on the Influences of Heredity and Environment

Resemblance of Siblings in Intelligence and Achievement, by Gertrude H. Hildreth. New York City: Teachers College, Columbia University, 1925. Pp. 65.

Three main points are covered in this piece of research (1) whether or not resemblance is greater or less between related than between unrelated individuals (2) the determination of the amount of the resemblance, (3) a study of the relative influence of environment and heredity in causing resemblance or difference which may be found.

A widely diversified amount of sibling data was collected from three distinct sources: a typical American town, represented largely by Anglo-American stock of average social, economic and intellectual status; a New York private school, drawing children from superior homes; and a Hebrew Orphan Asylum group whose wards come from inferior and even sordid environments. Amounts of resemblance were measured by correlation, total and partial, and the method of mean differences. Pairs of unrelated children taken at random were compared for resemblance with pairs of related children. To investigate the third point, above, four kinds of relationships were studied: (*a*) Resemblances of true sib orphans reared together. (*b*) Relation of unrelated orphans reared together for a longer or shorter period of

time. (*c*) Relation of unrelated orphans reared apart. (*d*) Resemblance of related orphans reared apart from a certain age, but eventually brought together at the asylum and tested.

Positive relation in intelligence and achievement was found for all pairs of true sibs. True sibs raised apart resemble each other about as much as true sibs reared together. There is no difference in the resemblance of unrelated children reared together and unrelated children reared apart, when selected at random from the general population. Heredity rather than environment seems to account for resemblances. GLADYS C. SCHWESINGER.

STATISTICAL ARGUMENTS FOR "GUIDANCE"

Conservation of Intelligence in Illinois High Schools, by Charles W. Odell. The University of Illinois, 1925. Pp. 55.

This bulletin presents data on the relation of intelligence to educational and vocational choices of over 1200 public high school seniors in Illinois. Comparison is made of scores in the Otis Self Administering Test of Mental Ability with "Information Blank" replies, to determine the waste of mental ability likely to result from the selection by these pupils of postgraduation careers not in accord with their intelligences or with the needs of society. Like other similar studies, cited and reviewed in the course of the discussion, this one is admittedly indeterminate because of a host of uncontrolled and immeasureable factors affecting the tests, the questionnaire answers, and the occupational-intelligence analysis. The data seem to show, however, that many pupils of indifferent mental ability are planning to go to college and prepare for the professions; that others of high intelligence are planning to go at once into unexacting occupations; and that the proportions selecting various vocations are in little relation to the present proportions in those vocations. About a third of the senior boys intend to take up engineering and about 60 per cent of the girls are looking toward teaching. No quantitative statement of the waste to be expected from poor educational and vocational choices is possible on the basis of this study. The principal net result of it, as it seems to the author himself, is the further accumulation of data supporting the demand for more educational and vocational guidance in the high school. M. H. WILLING.

An Attempt to Determine Differences between Neurotic and Normal Children

A Comparison of Neurotic and Normal Children in Public Schools, by Esther Katz Rosen. New York City: Teachers College, Columbia University, 1925. Pp. 51.

The purpose of this investigation was to compare the intellectual and educational status of a group of school children, diagnosed by a psychiatrist as neurotic, with an equal number of normal school children, keeping the factors of age, sex, race, and grade constant. Tests used include the Stanford-Binet, given twice to each child, the Pintner-Paterson Performance Scale, the Pintner Educational Survey, Handwriting and Cancellation Tests. The subjects studied were New York City or Baltimore children, all of Russian Jewish parentage. None of the performances studied revealed any differentiating responses from the two groups. It must be borne in mind, however, that the non-neurotic mates were as a group retarded in age-grade status and intelligence in order to match similar conditions in the neurotics. These latter, it was found, when compared with a total school population are retarded in age-grade status and are below par in intelligence. None of the test reactions offered data which might be of diagnostic value in objectively identifying the neurotic child. His handwriting is poor in quality, but up to standard in speed. His IQ is as constant after a period of two weeks as those reported by other investigators for groups of unselected children. Neurotics tend to scatter on the Stanford-Binet in much the same way as do normal children. These last two findings are interesting in that they tend to upset a belief to the contrary which has had some weight in the past. In all of her findings, Dr. Rosen concludes that the differences within one group are much greater than any differences to be found between the neurotic and non-neurotic groups. Gladys C. Schwesinger.

A Text for Beginning Students

Elementary Psychology, by Arthur I. Gates. New York: The Macmillan Co., 1925. Pp. 594.

This is a parallel volume to the author's *Psychology for Students of Education*, Macmillan, 1923. Both books cover pretty much the same ground, although they differ materially in content and applica-

tion. The new volume contains a more extensive account of the introspective studies of conscious states and treats at greater length the problems of learning, perception, imagination, and reasoning. The following quotations from the preface give an idea of the author's aim: "The book has been written with the limitations of the elementary student in mind. The method of development is of the spiral type . . . The book contains more practical applications than is usual in an introductory text." And again, "The results both of introspective and objective study have been utilized with equal freedom. The work is designed to be a survey rather than a system of psychology."

Since the text is for beginning students, no attempt has been made to make the survey cover all the fields of psychology. The book does not contain a chapter bearing any of the following titles; Attention, Will, Consciousness, Self, Association, Body-mind Relation, Mental Development, Industrial Psychology, Abnormal Psychology, Social Psychology, or Animal Psychology. Words which do not appear in the index include; behaviorism, Charcot, Fechner, fovea, gestalt, Herring, hypnotism, ideomotor, imitation, Janet, Jennings, Jung, Kraepelin, Lashley, Meumann, Müller-Lyer, nesting, nystagmus, paralysis, perseveration, Prince, purposive psychology, pyramidal, rotation, Royce, self, self-activity, Spearman, stereoscope, Stern, sublimation, syllogism, synesthesia, telegraphy, telepathy, thalamus, thirst. Although this list is to some extent misleading, since some of the content which the omitted words would suggest has been included under other headings, it does indicate that the author has purposely refrained from making the test an introduction to the entire science of psychology.

Compared to the usual type of introductory text in general psychology, this text is somewhat less systematic, less concerned with theoretical position or philosophical assumptions, inclined to avoid taking a very positive stand on controversial questions, and particularly strong on practical applications to education. If it is less suited than some other books as an introductory text for students who are to become professional psychologists, it is correspondingly well adapted to the needs of the general student, particularly the student of education. Indeed the book could very well be used as a text in educational psychology, and is probably the best available for this purpose when the course in educational psychology is not preceded by one in general psychology. The chapters on the rôle of dominant

urges in habit formation, the general laws of learning, economic learning, conditions that affect efficiency, intelligence, and the measurement of human traits, are especially good.

The author's style throughout is straightforward and clear without giving the impression of conscious striving for literary effect. The concreteness of content is indicated by the fact that the book contains 73 figures and very many tables. The questions and exercises at the end of each chapter are exceptionally helpful, and the titles in the numerous reference lists are well chosen.

LEWIS M. TERMAN.
Stanford University.

A NEW EDITION

How to Teach Silent Reading to Beginners, by Emma Watkins. Philadelphia: J. B. Lippincott Company, 1924. Pp. 142.

Some four years ago a first grade teacher of the University Elementary School at Iowa City prepared an account of her own classroom procedure in silent reading and a defense of her own theory. The book was introduced by Dr. Ernest Horn and seems to have been very well received in spite of rather obvious crudities and defects of emphasis to which this reviewer called attention when the first edition appeared. The second edition has been enlarged by the addition of four lessons but neither the publisher nor the author seem to have taken the pains to re-edit the book in the light of criticism. This is perhaps due to the fact that the book appeals to a more or less uncritical audience which is willing to overlook its peculiarities in view of the lessons, materials, and devices which it contains. The added lessons do not contribute materially to the practical value of the book.

LAURA ZIRBES.
Teachers College, Columbia University.